Journal of Medical Internet Research

Journal Impact Factor (JIF) (2023): 5.8
Volume 16 (2014), Issue 12 ISSN 1438-8871 Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Original Papers

Effectiveness of Different Web-Based Interventions to Prepare Co-Smokers of Cigarettes and Cannabis for Double Cessation: A Three-Arm Randomized Controlled Trial (e273)
Julia Becker, Severin Haug, Robin Sullivan, Michael Schaub. . . . 4

Dose-Response Effects of a Web-Based Physical Activity Program on Body Composition and Metabolic Health in Inactive Older Adults: Additional Analyses of a Randomized Controlled Trial (e265)
David Vroege, Carolien Wijsman, Karen Broekhuizen, Anton de Craen, Diana van Heemst, Frans van der Ouderaa, Willem van Mechelen, P Slagboom, Michael Catt, Rudi Westendorp, Evert Verhagen, Simon Mooijaart. . . . 20

Reducing Alcohol Use During Pregnancy Via Health Counseling by Midwives and Internet-Based Computer-Tailored Feedback: A Cluster Randomized Trial (e274)
Nickie van der Wulp, Ciska Hoving, Kim Eijmael, Math Candel, Wim van Dalen, Hein De Vries. . . . 32

Automated Indexing of Internet Stories for Health Behavior Change: Weight Loss Attitude Pilot Study (e285)
Ramesh Manuvinakurike, Wayne Velicer, Timothy Bickmore. . . . 46

Comparing Effects in Regular Practice of E-Communication and Web-Based Self-Management Support Among Breast Cancer Patients: Preliminary Results From a Randomized Controlled Trial (e295)
Elin Børøsund, Milada Cvancarova, Shirley Moore, Mirjam Ekstedt, Cornelia Ruland. . . . 72

How Doctors View and Use Social Media: A National Survey (e267)
James Brown, Christopher Ryan, Anthony Harris. . . . 102

Motivations for Contributing to Health-Related Articles on Wikipedia: An Interview Study (e260)
Nuša Fari , Henry Potts. . . . 111

Analysis of a Parent-Initiated Social Media Campaign for Hirschsprung's Disease (e288)
Kristy Wittmeier, Cindy Holland, Kendall Hobbs-Murison, Elizabeth Crawford, Chad Beauchamp, Brodie Milne, Melanie Morris, Richard Keijzer.
1 2 4

Therapeutic Affordances of Social Media: Emergent Themes From a Global Online Survey of People With Chronic Pain (e284)
Mark Merolli, Kathleen Gray, Fernando Martin-Sanchez. . . . 132

eMental Health Experiences and Expectations: A Survey of Youths' Web-Based Resource Preferences in Canada (e293)
Felicia Wetterlin, Marissa Mar, Erika Neilson, Gregory Werker, Michael Krausz. . . . 148

XSL•FO
RenderX

Cumulative Query Method for Influenza Surveillance Using Search Engine Data (e289)
Dong-Woo Seo, Min-Woo Jo, Chang Sohn, Soo-Yong Shin, JaeHo Lee, Maengsoo Yu, Won Kim, Kyoung Lim, Sang-Il Lee. 156

Use of Home Telemonitoring to Support Multidisciplinary Care of Heart Failure Patients in Finland: Randomized Controlled Trial (e282)
Anna-Leena Vuorinen, Juha Leppänen, Hannu Kaijanranta, Minna Kulju, Tiina Heliö, Mark van Gils, Jaakko Lähteenmäki. 172

Nonprobability Web Surveys to Measure Sexual Behaviors and Attitudes in the General Population: A Comparison With a Probability Sample Interview Survey (e276)
Bob Erens, Sarah Burkill, Mick Couper, Frederick Conrad, Soazig Clifton, Clare Tanton, Andrew Phelps, Jessica Datta, Catherine Mercer, Pam Sonnenberg, Philip Prah, Kirstin Mitchell, Kaye Wellings, Anne Johnson, Andrew Copas. 211

Recruitment via the Internet and Social Networking Sites: The 1989-1995 Cohort of the Australian Longitudinal Study on Women's Health (e279)
Gita Mishra, Richard Hockey, Jennifer Powers, Deborah Loxton, Leigh Tooth, Ingrid Rowlands, Julie Byles, Annette Dobson. 225

Use of an Electronic Patient Portal Among the Chronically Ill: An Observational Study (e275)
Iiris Riippa, Miika Linna, Ilona Rönkkö, Virpi Kröger. 235

Personal Health Record Reach in the Veterans Health Administration: A Cross-Sectional Analysis (e272)
Stephanie Shimada, Cynthia Brandt, Hua Feng, D McInnes, Sowmya Rao, James Rothendler, David Haggstrom, Erica Abel, Lisa Cioffari, Thomas Houston. 245

Consumer Participation in Quality Improvements for Chronic Disease Care: Development and Evaluation of an Interactive Patient-Centered Survey to Identify Preferred Service Initiatives (e292)
Elizabeth Fradgley, Christine Paul, Jamie Bryant, Ian Roos, Frans Henskens, David Paul. 258

Personal Experiences and Emotionality in Health-Related Knowledge Exchange in Internet Forums: A Randomized Controlled Field Experiment Comparing Responses to Facts Vs Personal Experiences (e277)
Joachim Kimmerle, Martina Bientzle, Ulrike Cress. 287

Which Frail Older Patients Use Online Health Communities and Why? A Mixed Methods Process Evaluation of Use of the Health and Welfare Portal (e278)
Peter Makai, Marieke Perry, Sarah Robben, Henk Schers, Maud Heinen, Marcel Olde Rikkert, René Melis. 294

Exploring the Relationship Between Changes in Weight and Utterances in an Online Weight Loss Forum: A Content and Correlational Analysis Study (e254)
Eric Hekler, Gaurav Dubey, David McDonald, Erika Poole, Victor Li, Elizabeth Eikey. 308

Impact of an Online Medical Internet Site on Knowledge and Practice of Health Care Providers: A Mixed Methods Study of the Spinal Cord Injury Rehabilitation Evidence Project (e296)
Janice Eng, Vanessa Noonan, Andrea Townson, Caroline Higgins, Jess Rogers, Dalton Wolfe. 328

Occupational Therapists' Views on Using a Virtual Reality Interior Design Application Within the Pre-Discharge Home Visit Process (e283)
Anita Atwal, Arthur Money, Michele Harvey. 339

A Search Engine to Access PubMed Monolingual Subsets: Proof of Concept and Evaluation in French (e271)
Nicolas Griffon, Matthieu Schuers, Lina Soualmia, Julien Grosjean, Gaétan Kerdelhué, Ivan Kergourlay, Badisse Dahamna, Stéfan Darmoni. .
3 5 3

XSL•FO
RenderX

Dr Google and the Consumer: A Qualitative Study Exploring the Navigational Needs and Online Health Information-Seeking Behaviors of Consumers With Chronic Health Conditions (e262)
Kenneth Lee, Kreshnik Hoti, Jeffery Hughes, Lynne Emmerton. . . . 364

The Effect of Individual Factors on Health Behaviors Among College Students: The Mediating Effects of eHealth Literacy (e287)
WanChen Hsu, ChiaHsun Chiang, ShuChing Yang. . . . 377

Movement-Based Interaction Applied to Physical Rehabilitation Therapies (e281)
Juan Garrido Navarro, Victor Ruiz Penichet, María Lozano Pérez. . . . 388

Reviews

Social Media and Mobile Apps for Health Promotion in Australian Indigenous Populations: Scoping Review (e280)
Carl Brusse, Karen Gardner, Daniel McAullay, Michelle Dowden. . . . 57

Interformat Reliability of Digital Psychiatric Self-Report Questionnaires: A Systematic Review (e268)
Sven Alfonsson, Pernilla Maathz, Timo Hursti. . . . 199

Educational Technologies in Problem-Based Learning in Health Sciences Education: A Systematic Review (e251)
Jun Jin, Susan Bridges. . . . 274

Viewpoint

Retaining Critical Therapeutic Elements of Behavioral Interventions Translated For Delivery via the Internet: Recommendations and an Example Using Pain Coping Skills Training (e245)
Christine Rini, Laura Porter, Tamara Somers, Daphne McKee, Francis Keefe. . . . 88

Tutorial

Web-Based Assessments of Physical Activity in Youth: Considerations for Design and Scale Calibration (e269)
Pedro Saint-Maurice, Gregory Welk. . . . 184

Short Paper

Ethical Issues in Using Twitter for Public Health Surveillance and Research: Developing a Taxonomy of Ethical Concepts From the Research Literature (e290)
Mike Conway. . . . 319

XSL•FO
RenderX

Original Paper

Effectiveness of Different Web-Based Interventions to Prepare Co-Smokers of Cigarettes and Cannabis for Double Cessation: A Three-Arm Randomized Controlled Trial

Julia Becker[1], MSc; Severin Haug[1], PhD; Robin Sullivan[1], MSc; Michael Patrick Schaub[1], PhD

Swiss Research Institute for Public Health and Addiction ISGF, University of Zurich, Zurich, Switzerland

Corresponding Author:
Julia Becker, MSc
Swiss Research Institute for Public Health and Addiction ISGF
University of Zurich
Konradstrasse 32
Zurich, 8031
Switzerland
Phone: 41 444411860
Fax: 41 444411870
Email: julia.becker@isgf.uzh.ch

Abstract

Background: The relationship between tobacco and cannabis use is strong. When co-smokers try to quit only one substance, this relationship often leads to a substitution effect, that is, the increased use of the remaining substance. Stopping the use of both substances simultaneously is therefore a reasonable strategy, but co-smokers rarely report feeling ready for simultaneous cessation. Thus, the question of how co-smokers can be motivated to attempt a simultaneous cessation has arisen. To reach as many co-smokers as possible, we developed brief Web-based interventions aimed at enhancing the readiness to simultaneously quit tobacco and cannabis use.

Objective: Our aim was to analyze the efficacy of three different Web-based interventions designed to enhance co-smokers' readiness to stop tobacco and cannabis use simultaneously.

Methods: Within a randomized trial, three brief Web-based and fully automated interventions were compared. The first intervention combined the assessment of cigarette dependence and problematic cannabis use with personalized, normative feedback. The second intervention was based on principles of motivational interviewing. As an active psychoeducational control group, the third intervention merely provided information on tobacco, cannabis, and the co-use of the two substances. The readiness to quit tobacco and cannabis simultaneously was measured before and after the intervention (both online) and 8 weeks later (online or over the phone). Secondary outcomes included the frequency of cigarette and cannabis use, as measured at baseline and after 8 weeks.

Results: A total of 2467 website users were assessed for eligibility based on their self-reported tobacco and cannabis co-use, and 325 participants were ultimately randomized and analyzed. For the post-intervention assessment, generalized estimating equations revealed a significant increase in the readiness to quit tobacco and cannabis in the total sample (B=.33, 95% CI 0.10-0.56, *P*=.006). However, this effect was not significant for the comparison between baseline and the 8-week follow-up assessment (*P*=.69). Furthermore, no differential effects between the interventions were found, nor were any significant intervention or time effects found on the frequency of tobacco or cannabis use.

Conclusions: In the new field of dual interventions for co-smokers of tobacco and cannabis, Web-based interventions can increase the short-term readiness to quit tobacco and cannabis simultaneously. The studied personalized techniques were no more effective than was psychoeducation. The analyzed brief interventions did not change the secondary outcomes, that is the frequency of tobacco and cannabis use.

Trial Registration: International Standard Randomized Controlled Trial Number (ISRCTN): 56326375; http://www.isrctn.com/ISRCTN56326375 (Archived by WebCite at http://www.webcitation.org/6UUWBh8u0).

(J Med Internet Res 2014;16(12):e273) doi:10.2196/jmir.3246

XSL•FO
RenderX

KEYWORDS

tobacco; cannabis; co-smoking; simultaneous cessation; motivational enhancement; personalized feedback; web-based intervention; motivational interviewing

Introduction

The Relationship Between Tobacco and Cannabis Use

Although smoking tobacco is the leading global cause of preventable death [1], cannabis is the most widely used illicit drug [2] and is associated with a range of physical and mental health problems [3,4]. Both substances are often used together; the majority of cannabis users also smoke cigarettes. In a study in the United States, 74% of marijuana users smoked cigarettes, compared to 29% of nonusers [5]. Furthermore, cannabis use is reportedly more common among cigarette smokers than it is among nonsmokers. In the National Survey on Drug Use and Health in the United States, the 30-day prevalence of cannabis use was 38% among tobacco smokers and only 11% among nonsmokers [6]. In a similar survey in Switzerland, cannabis use in the previous 12 months was reported by 28% of adolescents who smoked tobacco daily compared to 9% and 2% of the adolescents who were ex-smokers and never-smokers, respectively [7].

The mechanisms that link the use of both substances are assumed to go beyond the mechanisms that explain the co-use of drugs in general [8]. For instance, both substances are usually smoked (have a shared route of administration) and are often used simultaneously (co-administration), that is, tobacco is added to cannabis joints ("mulling") or is smoked directly after cannabis ("chasing") [8,9]. In Switzerland, 97% of cannabis users smoke cannabis joints mixed with tobacco [10].

In the context of cessation, the relationship between the substances is often problematic. For instance, tobacco smokers who also consume cannabis seem to make fewer efforts to quit tobacco [11] and tend to be less successful in quitting tobacco than tobacco-only smokers [12]. Furthermore, the cessation of one substance is frequently accompanied by an increased use of the other [13-15], and cessation programs that exclusively address tobacco appear to be less effective for co-smokers of cannabis [16,17].

Interventions for Tobacco and Cannabis Use

Despite these findings, interventions have typically targeted tobacco or cannabis use alone and have rarely addressed both substances simultaneously. One explanation for the separate treatments may be that in many industrialized countries, the treatment of cannabis dependence is provided by psychiatrists, whereas interventions for tobacco smokers are part of a more general public health system [18,19]. However, the body of literature on the relationship between tobacco and cannabis use is growing, and authors of recent reviews perceive a demand for double interventions that target tobacco and cannabis simultaneously [8,9,20]. In line with this notion, a preliminary study of the development of such a program has indeed revealed this demand [21]. The experts and the co-smokers who participated in the preliminary study considered a dual cessation intervention to be feasible.

However, the participants also expected only modest readiness to simultaneously quit tobacco and cannabis use; half of the surveyed co-smokers were unaware of the association between tobacco and cannabis use [21]. Due to this finding, the authors developed three brief online interventions to enhance co-smokers' awareness of the relationship between the substances and their readiness to simultaneously quit each substance. The purpose of the current study was to evaluate these online interventions and examine how co-smokers' readiness to simultaneously quit tobacco and cannabis can be augmented. Motivational constructs such as the stages of change and the readiness to quit have been shown to predict the subsequent engagement in interventions [22,23] and abstinence [24].

Because of its easy access and ubiquitous presence, the Internet arose as a potentially effective medium to reach a large number of co-smokers who might be unaware of the relationship between their tobacco use and their cannabis use. Personalized, normative feedback is one motivational technique that can be applied to Web-based interventions for substance use. Based on the social norms approach [25], such interventions typically include self-assessment sections and feedback sections in which the participants' behavior is compared to a reference sample. The overestimation of substance use by others is common and is positively associated with one's own use [26]. Web-based social norm interventions use this association and aim to correct the participants' erroneous perceptions. Primarily recruiting college students and targeting alcohol use, Web-based norms approaches for interventions have yielded promising results [27].

Another established technique for building motivation is motivational interviewing (MI), which uses a client-centered, directive counseling style to explore and reduce ambivalence and increase the intrinsic motivation for change [28]. Brief face-to-face interventions based on MI have been found to be effective in reducing cannabis use [29] and may assist in smoking cessation [30]. MI in Web-based interventions is usually applied as a chat-intervention but is not fully automated. However, the first promising results of fully automated MI have recently been revealed by a computer-based intervention that targets perinatal drug use [31].

For this study, we developed three Web-based interventions that apply the above-mentioned techniques, that is, normative feedback and MI. For an active control group, we used Web-based psychoeducation. In addition, to maintain the low threshold for Internet access and keep the study dropout rate as low as possible, the interventions were designed as brief single-session interventions.

Aims of the Study and Hypotheses

The main aim of this study was to evaluate the effectiveness of three Web-based interventions to enhance co-smokers' readiness to quit both tobacco and cannabis simultaneously. Our first hypothesis (H1) was that the tested interventions would be

XSL•FO
RenderX

effective in enhancing the readiness to simultaneously quit tobacco and cannabis use. Thus, we assumed a significant within-subjects effect for assessment time. Because particular interactive interventions that were tailored to individuals have shown promising effects in aiding smoking cessation [32], our second hypothesis (H2) was that interactive and tailored interventions, that is, an intervention based on MI and an intervention that provides normative feedback, would more effectively enhance co-smokers' readiness to quit tobacco and cannabis use simultaneously than would mere psychoeducation. Because MI has shown promising effects as a motivational enhancement strategy for cannabis users [33], we additionally hypothesized that this intervention would outperform the effectiveness of the normative feedback intervention (H3).

Furthermore, this study aimed to evaluate the three interventions as they pertained to secondary outcome variables, that is, the frequencies of tobacco and cannabis use. We had the same hypotheses for these outcomes as those explained above.

Methods

Study Design and Setting

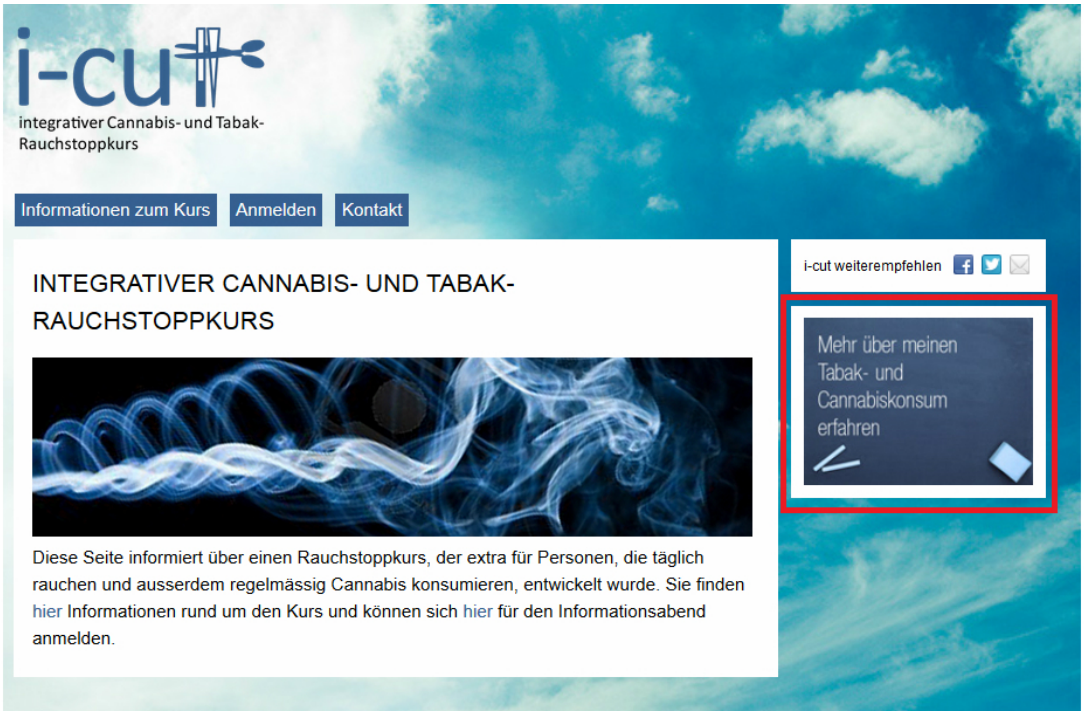
To test our hypotheses, we conducted a three-armed randomized trial with pre-, post- and 8-week follow-up assessments. (ISRCTN56326375; see Multimedia Appendix 1 [34] for the CONSORT EHEALTH checklist.)

The three single-session interventions were Web-based and fully automated. The baseline assessment (t0) was conducted at the beginning of the intervention session, and the post-intervention assessment (t1) immediately followed the intervention. After 8 weeks, the subjects were re-assessed (follow-up, t2) by phone or online. However, our primary focus was set on the immediate post-intervention assessment because our primary outcome, the readiness to simultaneously quit tobacco and cannabis, would not be applicable to participants who stopped smoking tobacco or cannabis after the intervention at t2. The data were collected between January and November 2012.

The interventions were integrated within the German-language website of the program "i-cut". In addition to the interventions, the website contained information about an integrative group cessation course for co-smokers of tobacco and cannabis. This cessation course is evaluated in a separate study (ISRCTN15248397).

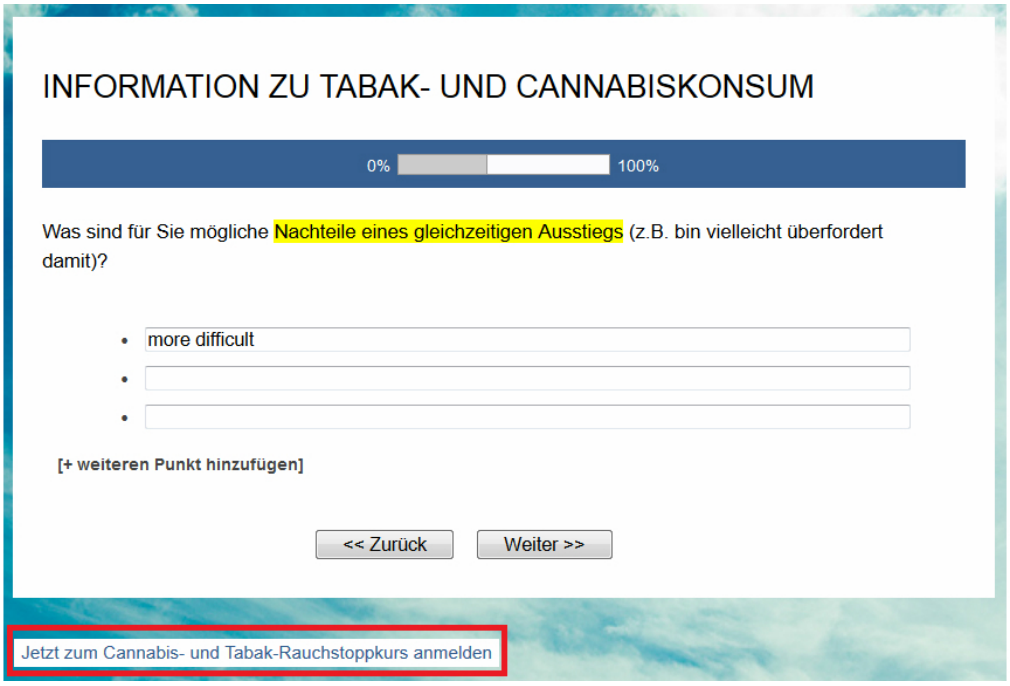
Participants could enter the current study in one of two ways. First, they could enter it directly from the start page of the website, where participants could choose between "getting more information about the course" and "learning more about my use of tobacco and cannabis". They were then directed to the course information pages or to the intervention session, respectively. We chose the cover term "learning more about my use of tobacco and cannabis" for the intervention session to attract co-smokers who were not seeking treatment. The second way to enter the study was to switch there from the course information pages by clicking a teaser that was displayed on the right side of each information page. It was also labeled "learning more about my use of tobacco and cannabis" (Figure 1). Conversely, participants could switch from the intervention session to the course information pages by clicking a hyperlink ("register now for the tobacco and cannabis cessation course"). This hyperlink was present on every page of the intervention, and the participants who clicked on it were directed to the course information pages and dropped out of the present study. Figure 2 shows a sample page of the intervention and the hyperlink.

Figure 1. Screenshot of the teaser (red square) for the Web-based intervention as displayed on the course information pages.



XSL•FO
RenderX

Figure 2. Screenshot of the intervention (intervention arm: motivational interviewing) and the hyperlink (red square) that directed participants to the Web pages with information about the smoking cessation course.



Eligibility and Recruitment

The inclusion criteria for study participation included any tobacco use during the past 4 weeks and any cannabis use during the past 6 months. As implicit inclusion criteria, participants had to speak German and be computer literate. There were no age restrictions or other exclusion criteria.

Recruitment for the present study ran parallel to recruitment for a feasibility study of the above-mentioned smoking cessation course for co-smokers of tobacco and cannabis. This was conducted in Zurich and a neighboring city. The recruitment strategy has been described in the publication on the course development [21]. Briefly stated, recruitment was carried out online and offline. First, a press release about the course was issued, which resulted in several reports in local newspapers and on radio and TV stations. Furthermore, brochures and pamphlets were sent to counseling centers for addiction prevention and treatment, psychiatrists, and health (care) centers in the canton of Zurich and in the bordering cantons. Additionally, two social media platforms and a teaser in the online edition of a popular free newspaper were used for recruitment. All of these referred readers to the start page of the website for more information.

To maximize the response rates, study participants were also offered the opportunity to participate in a lottery for three vouchers valued at 300, 200, or 100 Swiss Francs after they completed the first session, including the second measurement. Additionally, a second lottery for the same values served as an incentive to participate in the follow-up measurement. The data were collected from January to December 2012.

Procedure

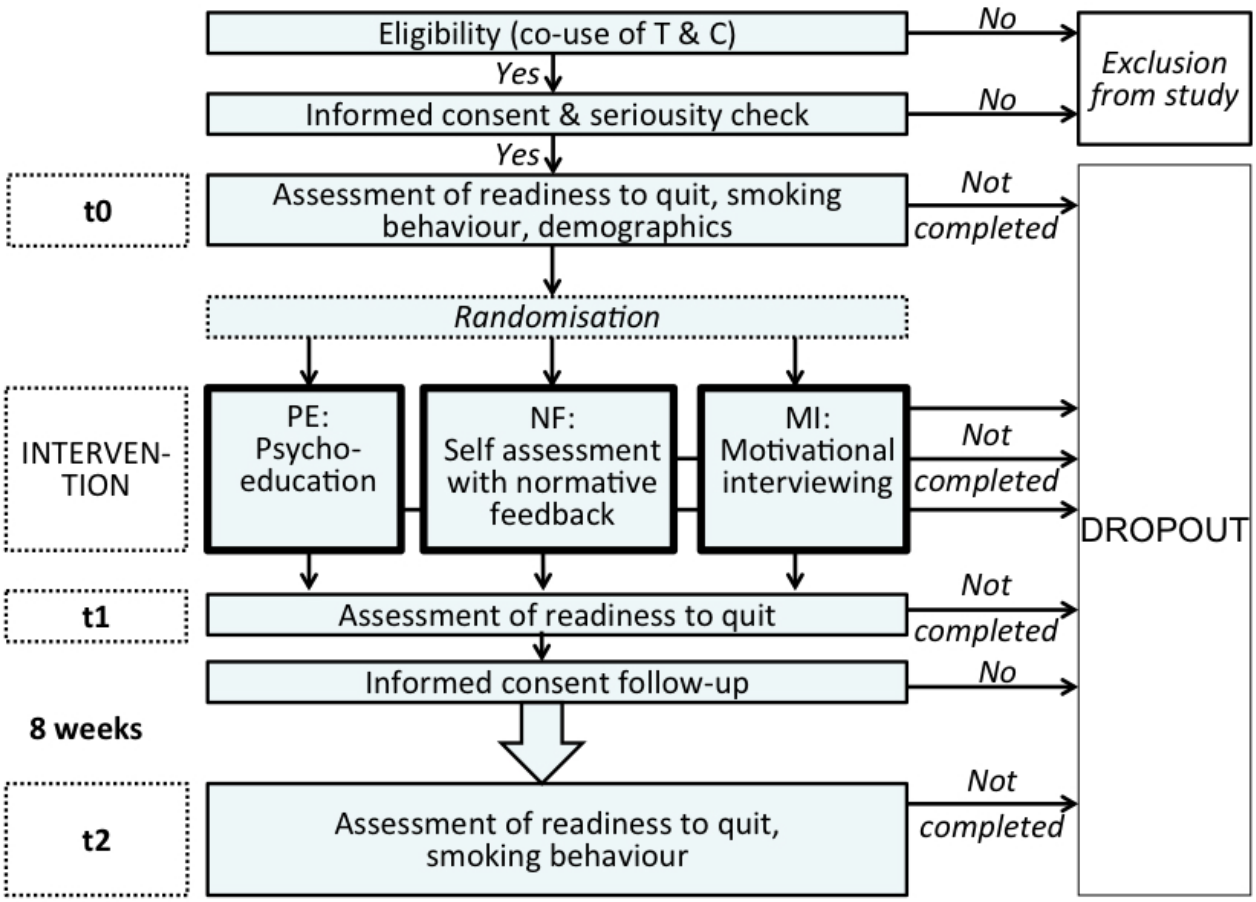
Figure 3 illustrates the study procedure in detail. The initial questions presented to potential participants were used to check the inclusion criteria. If the users met these criteria, they were informed about the opportunity to participate in a study that aimed to improve the website's information offerings. Those who did not meet the inclusion criteria or did not provide informed consent were excluded from the study and were referred to a webpage that contained information on tobacco and cannabis use, which was also provided in the psychoeducational intervention (see below). The study participants were then instructed to create an anonymous but personal identification code that combined certain letters of their parents' names and their own date of birth. The same procedure was applied at the follow-up assessment to link the data of the different assessments.

Once the baseline measurement (t0) was completed, participants were randomly assigned to one of three possible interventions. After finishing the intervention, participants were reassessed (post-intervention, t1) and informed about the 8-week follow-up assessment (t2). To keep the threshold for entering the study as low as possible, the information about the follow-up assessment was provided only at this point. This was done because the main aim of the present study was to enhance the readiness to simultaneously quit tobacco and cannabis between t0 and t1. Participants who provided informed consent for the follow-up assessment could indicate whether they wanted to answer the follow-up questionnaire online or over the phone. At the end of the session, participants were referred to the webpage of the group cessation program if they were interested.

For the follow-up assessment, participants were contacted after 8 weeks via their chosen medium (ie, via an email that included a link to the online questionnaire or via telephone). Those who preferred to answer the questionnaire online received an email reminder after approximately 2 weeks if they had not yet completed the online questionnaire. Those who chose the telephone questionnaire were contacted up to ten times.

XSL•FO
RenderX

Figure 3. Study procedure.



Interventions

General Information and Technological Background

Participation in the interventions was free, and access was open for every eligible participant. The delivery of interventions was fully automated. An open source software, LimeSurvey (Version 1.91), was used to program the survey and the interventions. As described below, the interventions varied in the extent to which they were interactive with the participants. Additionally, the interventions differed in the way in which each was tailored to the responses that the participants had given during the baseline assessment and during the interventions themselves.

Intervention 1: Normative Feedback

The first intervention contained a combination of self-assessment and personalized normative feedback (NF). It consisted of three sections that included one each for tobacco use, cannabis use, and co-smoking. In the first and second sections, participants began by completing a questionnaire: the Fagerstrom Test of Nicotine Dependence (FTND) [35,36] and the Cannabis Use Disorder Identification Test (CUDIT) [37], respectively. Participants received feedback following each questionnaire. Feedback was individually tailored to participants using an algorithm based on the results from the FTND, the CUDIT, and the baseline data. Based on the social norms approach, each participant's reported frequency of smoking was presented in relation to the normative data from Swiss community samples. Afterwards, participants received feedback about their questionnaire scores and were told whether their responses met the criteria for dependency (FTND) and/or problematic use (CUDIT), respectively. Explanations of "cigarette dependence" and "problematic cannabis use" were also given. Each substance-specific section concluded with brief recommendations for cessation or moderation of use. In addition, at the end of the intervention, information was provided that simultaneously accounted for the participant's use patterns of tobacco and cannabis. Participants who regularly smoked both tobacco and cannabis were informed about the group cessation course and referred to the end of the post-intervention assessment for further information. Participants who used either just one of the substances or both less regularly received contact details for the appropriate consulting services. Table 1 presents examples of translated feedback.

XSL•FO
RenderX

Table 1. Examples of feedback provided during the normative feedback intervention to a participant who smoked more than five cigarettes per day and used cannabis less than once per week.

Intervention step	Example
Feedback on tobacco use frequency	You indicated smoking an average of 12 cigarettes per day. Among Swiss males, 70% do not smoke at all. Only approximately 10% smoke more than you.
Feedback on cigarette dependence	Your nicotine dependence is classified as high. Your result means that quitting may be more difficult for you compared to people with low dependence. Presumably, you will experience withdrawal symptoms. Nevertheless, these symptoms will weaken soon, and there are helpful aids against them. For instance, nicotine replacement therapy is very effective. However, quitting smoking requires more than just getting through the withdrawal symptoms. For example, you should develop individual strategies to help you cope with risk situations where the temptation of smoking a cigarette is high. Professional support (eg, a smoking cessation course) can be very helpful in developing such strategies.
Feedback on cannabis use frequency	During the past 4 weeks, you used cannabis two or three times. A survey revealed that 89% of Swiss adolescents and young adults do not use cannabis at all. Only 4% use it more often than you.
Combined feedback	Of course, it is not easy to quit both substances simultaneously for good, especially after having smoked cigarettes on a regular basis. You can ask for support at [name of a center for addiction counseling and treatment] and mention that you also smoke joints occasionally.

Intervention 2: Motivational Interviewing

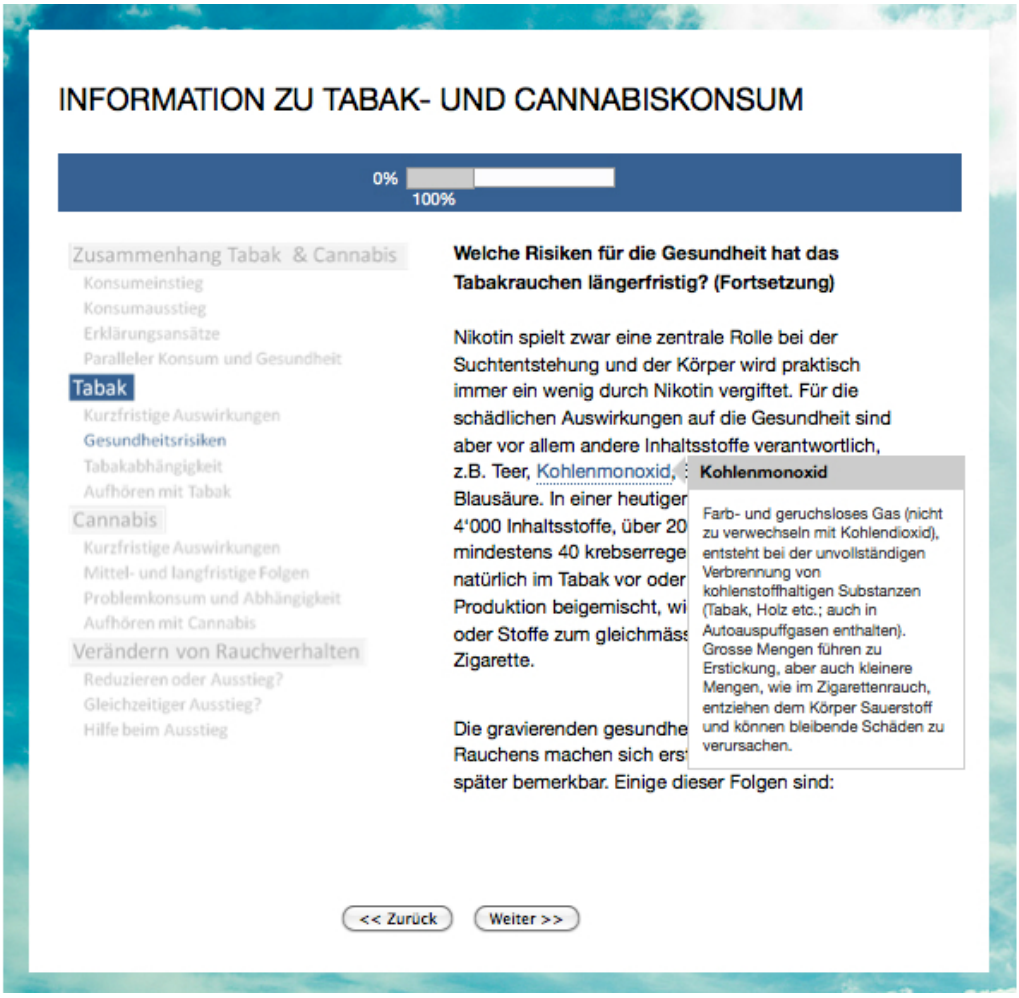
The second intervention was based on the principles of motivational interviewing (MI). It was highly interactive and tailored to the participant, and it used a selection of MI techniques that could be adapted to a Web-based intervention, such as open-ended questions, affirmative feedback, and periodic summaries. The aim of this intervention was to promote participants' self-reflective thinking about their own smoking behavior and intentions to change it and to enhance their self-confidence in the ability to change. This was done through different tasks, such as decisional balance tasks, in which participants wrote down personal pros and cons of stopping tobacco use, cannabis use, or both simultaneously (Figure 2). Participants were also asked to write down what advice they would give to a co-smoking friend and to indicate their confidence in successfully stopping tobacco, cannabis, or both simultaneously on a confidence ruler. Participants received feedback, including a brief summary of their indicated change in self-confidence and a brief informational text about the simultaneous cessation of tobacco and cannabis use. To further enhance their self-confidence, participants were asked to list any behavior that they had successfully changed in the past and to write down the names of persons in their network who could provide some level of social support during their attempt to quit smoking. Participants who, at baseline, had low levels of motivation to quit smoking and cannabis simultaneously received an additional task.

Intervention 3: Psychoeducation

The third intervention was the active control group. It provided psychoeducational information (PE) about tobacco and cannabis use. The information was thematically subdivided into smaller subsections. Participants had to read the sections in sequential order. Several terms and concepts that some readers may not know (eg, "carbon monoxide") were explained in a small text box that appeared when mousing over the word of interest (Figure 4). The PE intervention started with an explanation of the association between the two substances with regard to the initiation and cessation of their use, their linking mechanisms, and the potential health consequences of their co-use. The next chapter contained information about the short-term and long-term consequences of tobacco use, tobacco dependence, and the cessation of tobacco use and was followed by an analogous chapter on cannabis. The final chapter provided information about changing smoking behavior and addressed smoking reduction versus abstinence, the simultaneous cessation of tobacco and cannabis use, and support during the cessation process. At this point, the group cessation program was mentioned, and participants were referred to the end of the post-intervention assessment to receive further information.

XSL•FO
RenderX

Figure 4. Screenshot of a section in the psychoeducational intervention.



Outcome Measures

The primary outcome measure was participants' readiness to quit the use of tobacco and cannabis simultaneously. Readiness was measured at all three time points by the question, "To what extent are you ready to quit tobacco and cannabis simultaneously?" Participants indicated their readiness on a ruler ranging from 1 ("not at all") to 10 ("very much"). The item was designed based on the contemplation ladder [22], which is especially suited to measure the early stages of readiness. In addition, a comparison of the readiness ruler to other measures of motivation to change revealed its good concurrent and predictive validity and its superior clinical utility when its brevity and ease of administration are considered [38].

Secondary outcomes included the self-reported frequency of tobacco and cannabis use at baseline (t0) and at the 8-week follow-up (t2). The frequency of tobacco use was defined as the daily amount of cigarettes smoked during a typical smoking day, corrected for the number of smoking days during the past month. The frequency of cannabis use in the prior week was assessed using a 7-day timeline follow-back question [39].

Baseline Measures

At baseline, we assessed key demographic variables: age, sex, highest educational attainment, and nation of residence. Furthermore, we measured participants' smoking history: age of onset of tobacco and cannabis use and the number of prior attempts to quit tobacco use, cannabis use, or both simultaneously.

Statistical Analysis

According to the intention-to-treat principle, all participants who provided informed consent and communicated their intention to provide serious answers to the questionnaire were included in the analyses. Using the Amelia II multiple imputation package of the R software environment for statistical computing, Version 2.15.3 [40], we imputed 20 datasets. In a simulation study using data from an online self-help program for problem drinkers, Amelia II outperformed other methods of multiple imputation [41]. Hypotheses tests were performed using each dataset separately and were pooled afterwards (intention-to-treat analysis). To determine the robustness of our results to the analytic strategy, we also performed complete case analyses considering only participants who provided data at all three assessments.

The trial arm differences in the baseline measurements were tested using analysis of variance (ANOVA) and the Kruskal-Wallis test for continuous variables and chi-square tests for categorical variables, depending on each variable's parametric properties. We conducted dropout analyses for the post-intervention and the follow-up assessment using logistic regression analyses.

XSL•FO
RenderX

To analyze the primary and secondary outcome variables, we used generalized estimating equations (GEE) that considered the correlated nature of repeated measures. In the GEE models used to test H1, the only predictor was measurement time. The GEE models used to test differential effects of the interventions (H2 und H3) considered five variables: measurement time, intervention type, interaction of time and intervention type, gender, and baseline readiness to quit cannabis. Measurement time was entered as a dummy-coded categorical variable with the baseline measurement as the reference category. In the GEEs that modeled the secondary outcomes, time was binary because the frequency of use was measured only at t0 and t1. Interaction effects examined whether the intervention type had a differential effect on the changes in the outcome variable over time in the two compared groups, thus answering H2 and H3. Gender and baseline readiness to quit cannabis use were included as control variables because they differed between the two groups at baseline.

To directly test the postulated differential effects by intervention group, the interventions were grouped by hypothesis and analyzed in two separate models. One model contrasted the PE intervention as a reference category with the combination of MI and NF (H2), whereas the second model contrasted the two interactive, personalized intervention types, NF and MI, with each other (H3).

All GEE models were based on an unstructured working correlation matrix. For the models of readiness to quit, a normal model with an identity link function was chosen. In the models of frequency of tobacco and cannabis use, we used a negative binomial model with a log link function. In case of statistically significant results, Cohen's *d* was calculated. The alpha level was set at alpha=.05, and the analyses were performed using R [40], Stata 12.1 SE [42], IBM SPSS Statistics Version 20 [43], and G*Power 3.1.3 [44].

The power calculation was based on the outcome of primary interest, the readiness to simultaneously cease tobacco and cannabis use, as measured directly after the intervention (t1). The study was powered to detect a small effect because most reviews of Web-based interventions that aim at changing tobacco or cannabis use behavior report small intervention effects [45,46]. For a 3 intervention arms x 2 repeated measurements design, a total sample size of N=246 was required when assuming a small effect size of f=.10, according to Cohen [47], with a 2-sided, type I error rate alpha=.05, and a power of 80%.

Ethical Approval

The study was designed in accordance with the Declaration of Helsinki and was approved by the ethics committee of the Canton of Zurich, Switzerland (approval number: KEK-StV-Nr. 23/11, June 27, 2011, and amendment for the Internet-based intervention, November 11, 2011).

Results

Baseline Characteristics, Response Analysis, and Intervention Duration

As shown in Figure 5, 1631 of the 2476 users who were assessed for eligibility met inclusion criteria. Of those, less than a quarter (325/1631, 19.93%) provided informed consent and completed the baseline assessment and could therefore be randomized into one of the intervention groups. Of them, 80.3% (261/325) of participants completed the intervention and participated in the post-intervention assessment, and 26.2% (85/325) participated in the follow-up assessment.

Table 2 compares the baseline variables across the intervention groups. Except for the relatively high percentage of women (30.3% in the PE intervention vs 14.9% in the NF intervention and 17.6% in the MI intervention) and the higher readiness to quit cannabis use in the PE intervention, no baseline variables differed significantly across the three groups.

During the intervention and the post-intervention measurement (t1), 28.6% (93/325) of participants dropped out, and 28.3% (92/325) did not provide informed consent for the follow-up assessment. Of the 85 follow-up participants, 51 (60%) responded via the online questionnaire and 34 (40%) were followed up with via phone. The dropout analysis revealed that participation in the post-intervention assessment was predicted by the intervention condition and the readiness to stop tobacco and cannabis simultaneously at baseline. A lower readiness to stop simultaneously at baseline predicted a higher probability of participating in the post-intervention assessment: OR 0.89, 95% CI 0.80-0.988, SE 0.05, *P*=.03. Furthermore, participants in the NF condition had a participation rate of 94.7% and thus were significantly more likely to answer the post-intervention questions than the PE (73.4%) and the MI conditions (71.6%) (OR 6.32, CI 2.49-16.00, SE 0.47, *P*<.001). Participation at the follow-up assessment was predicted by sex, that is, males had a lower likelihood of participating at the follow-up than did females (OR 0.49, 95% CI 0.24-0.98, SE 0.36, *P*=.04). In contrast to the post-intervention assessment, the intervention groups did not significantly differ regarding their participation at the follow-up.

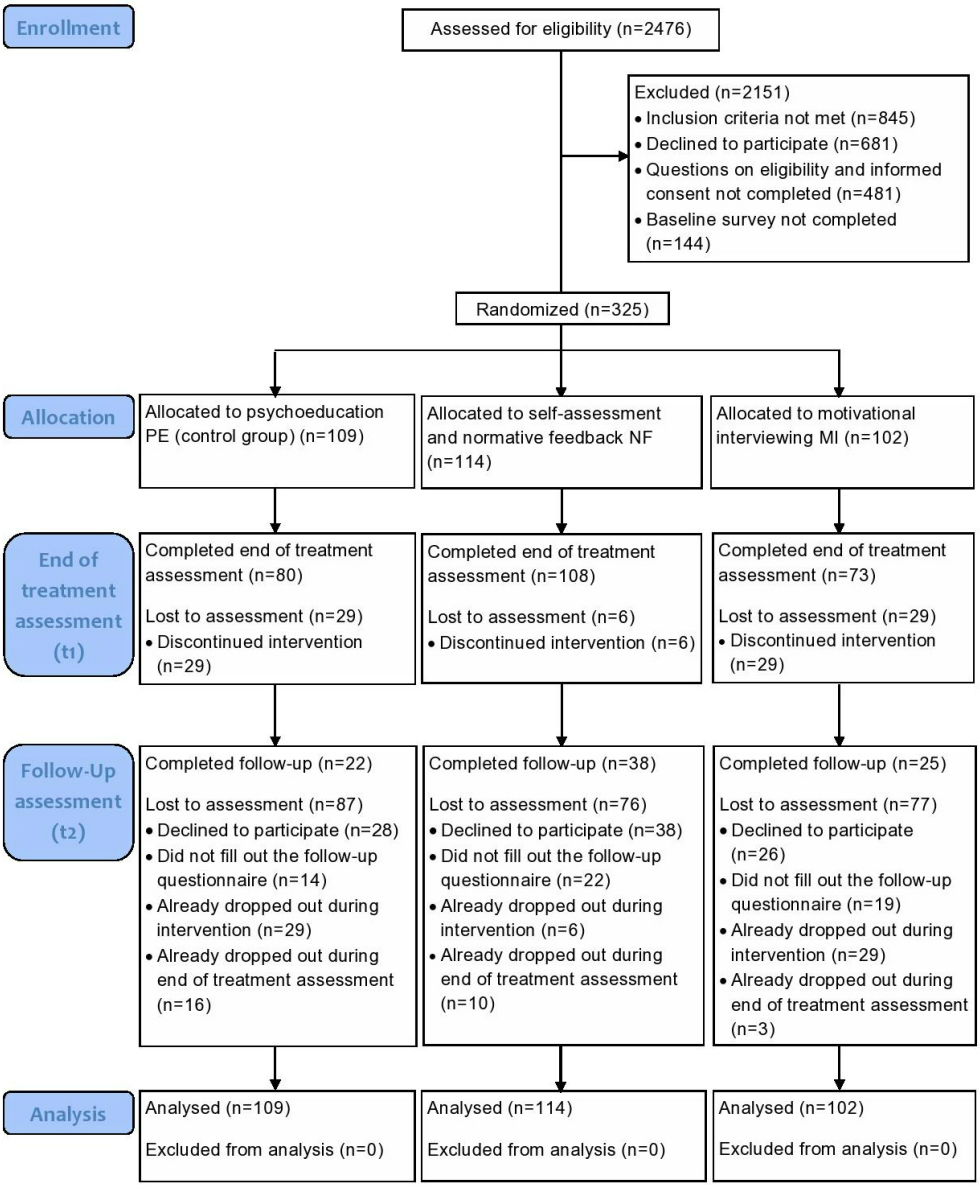
The duration of the interventions differed significantly. Overall, participants remained in the intervention sessions for an average of 25.5 minutes (SD 33.0), including the baseline and post-intervention assessments. Whereas the participants in the NF condition finished the session after a mean of 17.0 minutes (SD 9.1) on average, participants in the PE (mean 28.4, SD 38.4) and the MI (mean 28.9, SD 41.6) interventions stayed significantly longer ($F_{2,322}$=4.7, *P*=.01). However, there were no significant associations between the intervention duration and the outcomes, that is, the readiness to stop tobacco and cannabis simultaneously at the post-intervention assessment (*r*=.08, *P*=.22), the readiness to stop tobacco and cannabis simultaneously at the follow-up (*r*=.06, *P*=.60), tobacco use frequency at the follow-up (*r*=.15, *P*=.20), or cannabis use frequency at the follow-up (*r*=-.08, *P*=.49).

XSL•FO
RenderX

Table 2. Trial arm differences in baseline variables.

	PE (n=109)	NF (n=114)	MI (n=102)	Significance	
				F(df) / χ^2(df)	P
Females, n (%)	33 (30.3)	17 (14.9)	18 (17.6)	8.91 (2)	.01
Age in years, mean (SD)	30.5 (9.5)	29.2 (9.6)	29.6 (9.5)	1.22 (2)	.54
Tobacco use frequency (cigarettes per day), mean (SD)	12.5 (7.7)	12.0 (8.2)	13.6 (8.6)	2.16 (2)	.34
Cannabis use frequency (times per day), mean (SD)	2.5 (1.9)	2.3 (2.3)	2.3 (2.2)	3.54 (2)	.17
Age of tobacco use onset in years, mean (SD)	16.0 (3.2)	15.8 (2.9)	16.0 (2.7)	0.17 (2,322)	.84
Age of cannabis use onset in years, mean (SD)	17.1 (4.4)	16.3 (3.4)	16.5 (3.1)	1.09 (2,322)	.34
Prior simultaneous cessation attempt, n (%)	32 (29.4)	38 (33.6)	32 (31.4)	0.47 (2)	.79
Readiness to quit tobacco, mean (SD)	7.2 (2.4)	7.0 (2.7)	7.5 (2.4)	1.42 (2)	.49
Readiness to quit cannabis, mean (SD)	5.8 (3.0)	4.8 (3.3)	5.1 (2.9)	6.03 (2)	.049
Readiness to quit tobacco and cannabis simultaneously, mean (SD)	5.2 (2.8)	4.7 (3.0)	5.0 (2.9)	2.40 (2)	.30

Figure 5. Study flow chart.



XSL•FO
RenderX

Effects of the Intervention on Readiness to Simultaneously Quit Tobacco and Cannabis Use

As shown in Figure 6, the readiness to quit tobacco and cannabis use simultaneously slightly increased in all interventions between t0 and t1 and decreased thereafter. The GEE analysis used to test time effects in the total sample (H1) revealed that readiness to simultaneously quit was significantly higher at post-intervention than at baseline (B=.33, SE 0.12, 95% CI 0.10-0.56, *P*=.006). The effect size for the mean difference between these two assessments was small (*d*=0.20). At the follow-up assessment, the readiness to simultaneously quit was no longer significantly higher than at the baseline level (B=-.13, SE 0.33, 95% CI -0.81-0.54, *P*=.69).

Table 3 displays the results of the GEE models that tested the two hypotheses related to the differential change in readiness to quit tobacco and cannabis use simultaneously. Regarding H2, there was a significant main effect of time in examining the change in readiness to quit simultaneously from baseline (t0) to the post-intervention (t1) assessment. The effect size for this time effect, now based on the reference group PE, was small (*d*=0.38). This effect was not maintained at follow-up (t2). Furthermore, neither the intervention effect nor the time*intervention interaction was significant. As the analysis of H3 revealed, there were no significant time effects of readiness to quit simultaneously when only MI and NF were included in the model. In both models, the control variable baseline readiness to stop cannabis use at baseline was a significant predictor of readiness to stop both tobacco and cannabis use simultaneously.

The complete case analyses replicated these findings. The first model, which used the total sample, revealed a significant time effect at t1 (B=.31, SE 0.11, 95% CI 0.37-1.54, *P*=.001) but not at t2 (*P*=.17). Regarding H2, only the time effect observed when comparing the post-intervention with the baseline assessment was significant (B=.95, SE 0.18, 95% CI 0.37-1.54, *P*=.001). In the model used to test H3, there was no significant effect, except for the control variable baseline readiness to quit cannabis use.

Table 3. Results from the linear GEE models (with 20 imputed datasets) that examined readiness to quit tobacco and cannabis use simultaneously, according to H2 and H3.

Hypothesis	Parameter	B	Standard error	95% CI		*P*
				lower	upper	
H2	Intercept	1.83	0.38	1.07	2.58	<.001
	Groups NF & MI[a]	0.07	0.25	-0.43	0.56	.80
	Time t2[b]	-0.06	0.42	-0.89	0.78	.90
	Time t1[b]	0.59	0.19	0.22	0.96	.002
	Time t2[b]× Groups NF & MI[a]	-0.12	0.40	-0.90	0.66	.76
	Time t1[b]× Groups NF & MI[a]	-0.40	0.22	-0.83	0.03	.07
	Baseline readiness to stop cannabis use	0.57	0.04	0.48	0.65	<.001
	Female gender[c]	0.17	0.29	-0.40	0.74	.56
H3	Intercept	1.59	0.45	0.70	2.47	<.001
	Group MI[d]	0.20	0.29	-0.37	0.76	.50
	Time t2[b]	-0.03	0.42	-0.88	0.81	.94
	Time t1[b]	0.22	0.14	-0.06	0.50	.13
	Time t2[b]× Group MI[d]	-0.30	0.45	-1.18	0.58	.50
	Time t1[b]× Group MI[d]	-0.06	0.28	-0.60	0.48	.83
	Baseline readiness to stop cannabis use	0.59	0.05	0.49	0.69	<.001
	Female gender[c]	0.29	0.41	-0.50	1.09	.47

[a]reference: PE.

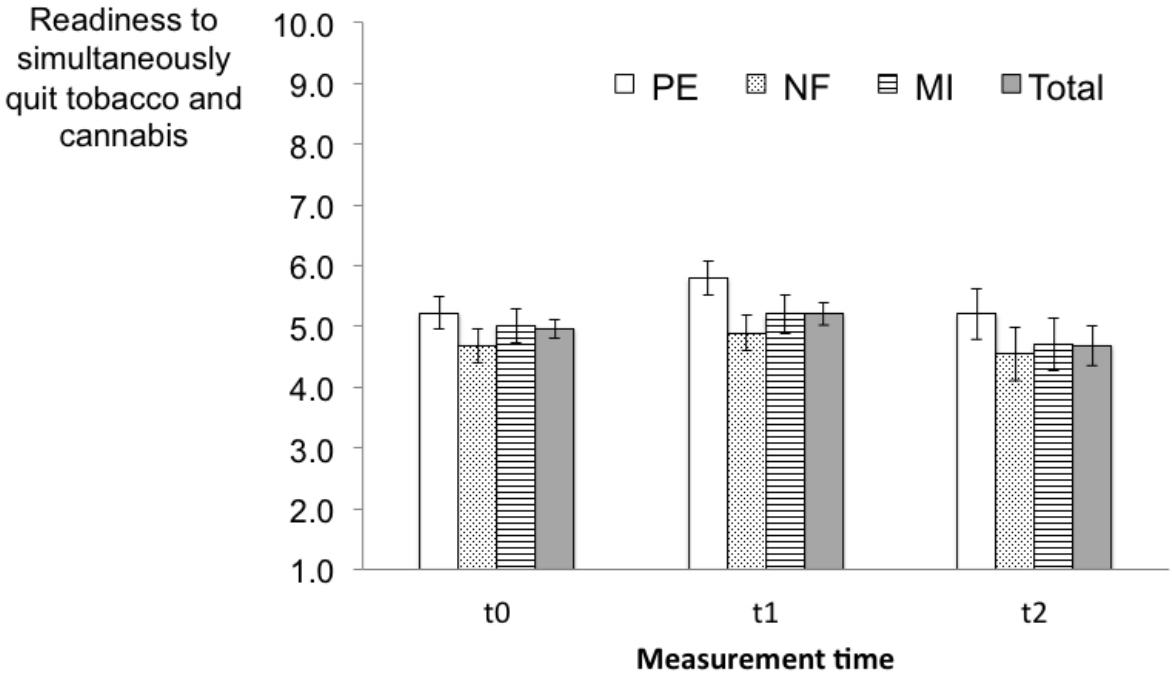
[b]reference: Time t0 (baseline).

[c]reference: male gender.

[d]reference: NF.

XSL•FO
RenderX

Figure 6. Course of readiness to simultaneously quit tobacco and cannabis use over time, pooled means of the intention-to-treat sample, including imputations; error bars represent the standard error of the mean (PE=psychoeducation, NF=normative feedback, MI=motivational interviewing).



Effects of the Intervention on Secondary Outcomes

Descriptive statistics for the frequency of tobacco and cannabis use showed only a weak decrease in frequency between the baseline and follow-up assessments (Table 4). The GEE model that analyzed H1 did not reveal a significant time effect for the frequency of either tobacco use (Incidence Risk Ratio [IRR] -0.04, SE 0.09, 95% CI -0.23-0.15, *P*=.70) or cannabis use (IRR -0.05, SE 0.12, 95% CI -0.29-0.20, *P*=.70).

None of the analyses of differential changes in the frequency of either tobacco or cannabis use revealed a significant time effect, intervention effect or time*intervention interaction.

The complete case analyses predominantly replicated these findings, revealing no significant time, group, or time*group interaction effects in the GEE models of tobacco or cannabis use frequency. One exception was a significant time effect for the frequency of tobacco use among the total sample (H1: IRR 0.89, SE 0.04, 95% CI 0.81-0.98, *P*=.02).

Table 4. Means and standard deviations of the frequency of tobacco and cannabis use at baseline (t0) and 8-week (t2) follow-up (descriptive statistics were calculated using the 20 imputed datasets).

Outcome variable	Time points	PE, mean (SD)	NF, mean (SD)	MI, mean (SD)	Total
Tobacco use frequency, cigarettes per day	t0	12.5 (2.4)	12.0 (2.5)	13.6 (2.5)	12.7 (2.5)
	t2	12.5 (2.5)	11.0 (2.5)	13.4 (2.7)	12.3 (2.6)
Cannabis use frequency, times per week	t0	2.5 (1.4)	2.3 (1.4)	2.3 (1.4)	2.3 (1.4)
	t2	2.4 (1.4)	2.2 (1.4)	2.2 (1.4)	2.2 (1.4)

Discussion

Principal Results

This study evaluated three brief fully automated Web-based interventions that aimed to enhance co-smokers' readiness to simultaneously quit their tobacco and cannabis use. Regarding the readiness to simultaneously quit using tobacco and cannabis, we assumed that all participants would have an increased level of readiness after the intervention compared to the baseline assessment (H1). Furthermore, we hypothesized that more interactive and individually tailored interventions would be more effective than mere information provision (psychoeducation; H2). Additionally, we tested the hypothesis that a Web-based intervention that applies principles derived from MI would be even more effective than an intervention that provides tailored, normative feedback (H3). The hypotheses regarding tobacco and cannabis use frequency were analogous.

Regarding the readiness to simultaneously quit tobacco and cannabis use, the results supported our first hypothesis. That is, in the total sample, the readiness to simultaneously quit was significantly elevated at the post-intervention assessment relative to baseline. This effect had disappeared by the 8-week follow-up assessment. The two hypotheses that assumed differential intervention effects were also not supported. With regard to the

XSL•FO
RenderX

frequency of tobacco and cannabis use, our analyses did not reveal time or intervention effects.

Strengths and Limitations

Among the strengths of this study is that it is the first study of Web-based interventions that target co-smokers of tobacco and cannabis. Furthermore, the interventions are fully automated and therefore require no personnel beyond their initial development.

Among the limitations of this study was its high attrition rate regarding participation at the follow-up assessment. However, high attrition rates are common in eHealth studies and brief interventions [48]. We addressed this limitation by using multiple imputation methods and performing traditional complete case analyses. Furthermore, our primary focus was set on the post-intervention assessment, which had a lower attrition rate. However, the dropout analyses revealed that participants dropped out selectively before the post-intervention assessment; that is, participants of the NF-condition were more likely to complete the post-intervention assessment, indicating the attrition superiority of normative feedback to the other two conditions. More difficult to explain is the finding that participants with a lower baseline readiness to quit simultaneously were more likely to complete the post-intervention assessment. A possible explanation of this finding might be that lower baseline readiness to quit simultaneously resulted in more uncertainty when confronted with the interventions. Possibly, this uncertainty resulted in a greater interest in the interventions and made participants following the intervention for a longer time than those participants that were already more ready for simultaneous cessation. Regarding the follow-up assessment, there was only one significant predictor of participation: female sex. However, this variable was statistically controlled in the GEE analyses because the proportion of males and females also differed between the intervention conditions.

A further limitation is that we did not include an assessment-only control group and could therefore not control for baseline assessment effects. We did not include such a control group because the baseline assessment, the intervention, and the post-intervention assessment happened in the same session. Therefore, individuals in an assessment-only control group would have been reassessed after less than 30 minutes.

Finally, the NF and MI interventions differed in length. The intervention sessions for MI and PE participants both lasted nearly 30 minutes, but the duration of an NF intervention session was approximately 11 minutes shorter. This difference was also reflected in the higher participation rate among NF participants in the post-intervention session. The possibility that we would have achieved significant differences between these two interventions if they had been equally long can therefore not be excluded.

Comparison With Prior Findings

The comparability to prior studies is limited because no Web-based interventions that target the co-use of tobacco and cannabis have been published. Additionally, Web-based MI interventions that are delivered fully automated and do not use chat-based MI-counseling are rare. However, the significant time effect in our study and the absence of differential intervention effects on readiness to quit are in line with the findings of a study that compared a single-session of MI-based chat-intervention with a chat in which participants received technical information about the baseline self-test [49]. That study included problematic alcohol and cannabis users but targeted only the particular problem behavior. The interventions were comparable to the MI and PE interventions of the current study with regard to their length but differed from the current interventions because they were not delivered in an automated fashion.

Moreover, we speculate that providing knowledge was a relatively effective measure in our study because co-smokers' baseline knowledge about the relationship between tobacco and cannabis use seems to be generally modest [21]. In addition, the psychoeducational intervention was the only intervention that provided information on the risk of physical harm from cannabis use. In one previous study, awareness of this risk was a significant predictor of readiness to simultaneously quit [21].

There are several possible explanations for the lack of time effects on the frequency of tobacco and cannabis use. First, the interventions were conceptualized as motivational enhancement interventions and targeted co-smokers who were in earlier stages in the process of behavior change. The interventions therefore had mainly motivational contents and only very few elements that are commonly applied to support the cessation or reduction of tobacco or cannabis use, such as the development of personal strategies or skills training. It has been previously shown that the effectiveness of Internet interventions in creating behavior change is associated with the incorporation of behavior change techniques [50]. In addition, the studied interventions were very brief compared to Web-based treatment interventions, which revealed significant effects on either tobacco use [51,52] or cannabis use [53,54]. This explanation is supported by the fact that other studies that analyzed Web-based interventions of a comparable length also revealed no effect on cannabis use [49,55]. It should also be considered that the current interventions targeted two behaviors simultaneously, which may require especially intensive interventions. Finally, the high attrition rate limits the interpretability of our findings concerning behavior change. This limitation is also illustrated by the different findings from the intent-to-treat and the complete case analysis.

Moreover, the appropriateness of fully automated MI is questionable because some components of the MI approach, such as therapeutic rapport, cannot be realized in an online setting. The efficacy of fully automated MI might be particularly limited when two behaviors are targeted simultaneously because dual cessation presumably provokes ambivalence that cannot be counterbalanced by a therapist. However, in brief face-to-face interventions for universal drug prevention and early intervention, MI was also not more effective than advice [56,57]. Compared to our study, however, significant changes over time in tobacco, cannabis, and alcohol use were achieved in both intervention groups. Furthermore, a Web-based intervention to promote smoking cessation using seven 45-minute sessions with MI-based video-chat in a virtual reality world revealed

XSL•FO
RenderX

both significant time and intervention effects but used an assessment-only control condition [51].

Conclusions

The findings of this study suggest that brief fully automated Web-based interventions have a short-term but perhaps no longer-term effect on co-smokers' readiness to simultaneously quit tobacco and cannabis use. There were no differential intervention effects, indicating that psychoeducation is not less effective compared to more individualized, interactive interventions when the co-use of tobacco and cannabis is targeted. Moreover, neither time nor intervention effects on substance use behavior were found. For dual-health behavior change, more intensive interventions regarding the length and the mode of administration (fully automated vs face-to-face, text-chat, or video-chat) may be needed. Future studies could examine more comprehensive Web-based treatment interventions for co-smokers and examine the efficacy of chat-based MI-counseling in this target group.

Acknowledgments

We would like to thank Fiona Herzig for her assistance in data collection and all participants of the study.

The Swiss Tobacco Control Fund provided financial support for this study (grant number 11.002932). The funding institution had no role in the development or evaluation of the interventions.

Conflicts of Interest

The authors were involved in the development of the interventions.

Multimedia Appendix 1

CONSORT EHEALTH checklist V1.6.1 [57].

[PDF File (Adobe PDF File), 994KB - jmir_v16i12e273_app1.pdf]

References

1. World Health Organization. WHO report on the global tobacco epidemic: warning about the dangers of tobacco. 2011. URL: http://whqlibdoc.who.int/publications/2011/9789240687813_eng.pdf [accessed 2014-01-13] [WebCite Cache ID 6Mb0hcQ69]
2. UNODC. World drug report. 2012. URL: http://www.unodc.org/documents/data-and-analysis/WDR2012/WDR_2012_web_small.pdf [accessed 2014-01-13] [WebCite Cache ID 6MazvijNn]
3. Caldeira KM, O'Grady KE, Vincent KB, Arria AM. Marijuana use trajectories during the post-college transition: health outcomes in young adulthood. Drug Alcohol Depend 2012 Oct 1;125(3):267-275 [FREE Full text] [doi: 10.1016/j.drugalcdep.2012.02.022] [Medline: 22464050]
4. Hall W, Degenhardt L. Adverse health effects of non-medical cannabis use. Lancet 2009 Oct 17;374(9698):1383-1391. [doi: 10.1016/S0140-6736(09)61037-0] [Medline: 19837255]
5. Richter KP, Kaur H, Resnicow K, Nazir N, Mosier MC, Ahluwalia JS. Cigarette smoking among marijuana users in the United States. Subst Abus 2004 Jun;25(2):35-43. [Medline: 15982966]
6. Substance Abuse And Mental Health Services Administration. National Survey on Drug Use and Health 2013. 2014. Table 6.9B – Types of Illicit Drug Use in the Past Month among Persons Aged 18 to 25, by Past Month Cigarette Use: Percentages, 2012 and 2013 URL: http://www.samhsa.gov/data/population-data-nsduh/reports [accessed 2014-11-15]
7. Radtke T, Keller R, Krebs H, Hornung R. Der Tabakkonsum Jugendlicher und junger Erwachsener in den Jahren bis 2001 bis 2009/10. 2011. URL: http://www.tabakmonitoring.ch/Berichte/Jugendliche/Forschungsbericht/Gesamt_jugend_10_dt.pdf [accessed 2014-11-14] [WebCite Cache ID 6JcipNrvN]
8. Agrawal A, Budney AJ, Lynskey MT. The co-occurring use and misuse of cannabis and tobacco: a review. Addiction 2012 Jul;107(7):1221-1233 [FREE Full text] [doi: 10.1111/j.1360-0443.2012.03837.x] [Medline: 22300456]
9. Peters EN, Budney AJ, Carroll KM. Clinical correlates of co-occurring cannabis and tobacco use: a systematic review. Addiction 2012 Aug;107(8):1404-1417 [FREE Full text] [doi: 10.1111/j.1360-0443.2012.03843.x] [Medline: 22340422]
10. Baggio S, Deline S, Studer J, Mohler-Kuo M, Daeppen JB, Gmel G. Routes of administration of cannabis used for nonmedical purposes and associations with patterns of drug use. J Adolesc Health 2014 Feb;54(2):235-240. [doi: 10.1016/j.jadohealth.2013.08.013] [Medline: 24119417]
11. Ford DE, Vu HT, Anthony JC. Marijuana use and cessation of tobacco smoking in adults from a community sample. Drug Alcohol Depend 2002 Aug 1;67(3):243-248. [Medline: 12127195]
12. Abrantes AM, Lee CS, MacPherson L, Strong DR, Borrelli B, Brown RA. Health risk behaviors in relation to making a smoking quit attempt among adolescents. J Behav Med 2009 Apr;32(2):142-149 [FREE Full text] [doi: 10.1007/s10865-008-9184-1] [Medline: 19003525]

XSL•FO
RenderX

13. Akre C, Michaud PA, Berchtold A, Suris JC. Cannabis and tobacco use: where are the boundaries? A qualitative study on cannabis consumption modes among adolescents. Health Educ Res 2010 Feb;25(1):74-82 [FREE Full text] [doi: 10.1093/her/cyp027] [Medline: 19515745]
14. Copersino ML, Boyd SJ, Tashkin DP, Huestis MA, Heishman SJ, Dermand JC, et al. Quitting among non-treatment-seeking marijuana users: reasons and changes in other substance use. Am J Addict 2006;15(4):297-302. [doi: 10.1080/10550490600754341] [Medline: 16867925]
15. Amos A, Wiltshire S, Bostock Y, Haw S, McNeill A. 'You can't go without a fag...you need it for your hash'--a qualitative exploration of smoking, cannabis and young people. Addiction 2004 Jan;99(1):77-81. [Medline: 14678065]
16. Gourlay SG, Forbes A, Marriner T, Pethica D, McNeil JJ. Prospective study of factors predicting outcome of transdermal nicotine treatment in smoking cessation. BMJ 1994 Oct 1;309(6958):842-846 [FREE Full text] [Medline: 7950614]
17. Moore BA, Budney AJ. Tobacco smoking in marijuana-dependent outpatients. J Subst Abuse 2001;13(4):583-596. [Medline: 11775084]
18. Hall SM, Prochaska JJ. Treatment of smokers with co-occurring disorders: emphasis on integration in mental health and addiction treatment settings. Annu Rev Clin Psychol 2009;5:409-431 [FREE Full text] [doi: 10.1146/annurev.clinpsy.032408.153614] [Medline: 19327035]
19. Prochaska JJ. Failure to treat tobacco use in mental health and addiction treatment settings: a form of harm reduction? Drug Alcohol Depend 2010 Aug 1;110(3):177-182 [FREE Full text] [doi: 10.1016/j.drugalcdep.2010.03.002] [Medline: 20378281]
20. Ramo DE, Liu H, Prochaska JJ. Tobacco and marijuana use among adolescents and young adults: a systematic review of their co-use. Clin Psychol Rev 2012 Mar;32(2):105-121 [FREE Full text] [doi: 10.1016/j.cpr.2011.12.002] [Medline: 22245559]
21. Becker J, Hungerbuehler I, Berg O, Szamrovicz M, Haubensack A, Kormann A, et al. Development of an integrative cessation program for co-smokers of cigarettes and cannabis: demand analysis, program description, and acceptability. Subst Abuse Treat Prev Policy 2013;8:33 [FREE Full text] [doi: 10.1186/1747-597X-8-33] [Medline: 24025478]
22. 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]
23. Hogue A, Dauber S, Morgenstern J. Validation of a contemplation ladder in an adult substance use disorder sample. Psychol Addict Behav 2010 Mar;24(1):137-144 [FREE Full text] [doi: 10.1037/a0017895] [Medline: 20307121]
24. Velicer WF, Redding CA, Sun X, Prochaska JO. Demographic variables, smoking variables, and outcome across five studies. Health Psychol 2007 May;26(3):278-287. [doi: 10.1037/0278-6133.26.3.278] [Medline: 17500614]
25. Perkins HW. The social norms approach to preventing school and college age substance abuse: a handbook for educators, counselors, and clinicians. San Francisco: Jossey-Bass; 2003.
26. Bertholet N, Faouzi M, Studer J, Daeppen JB, Gmel G. Perception of tobacco, cannabis, and alcohol use of others is associated with one's own use. Addict Sci Clin Pract 2013;8:15 [FREE Full text] [doi: 10.1186/1940-0640-8-15] [Medline: 24499600]
27. Moreira MT, Smith LA, Foxcroft D. Social norms interventions to reduce alcohol misuse in university or college students. Cochrane Database Syst Rev 2009(3):CD006748. [doi: 10.1002/14651858.CD006748.pub2] [Medline: 19588402]
28. Miller WR, Rollnick S, Conforti K. Motivational Interviewing: Preparing People for Change, 2nd edition. New York: The Guilford Press; 2002.
29. Stephens RS, Roffman RA, Curtin L. Comparison of extended versus brief treatments for marijuana use. J Consult Clin Psychol 2000 Oct;68(5):898-908. [Medline: 11068976]
30. Lai DT, Cahill K, Qin Y, Tang JL. Motivational interviewing for smoking cessation. Cochrane Database Syst Rev 2010(1):CD006936. [doi: 10.1002/14651858.CD006936.pub2] [Medline: 20091612]
31. Ondersma SJ, Chase SK, Svikis DS, Schuster CR. Computer-based brief motivational intervention for perinatal drug use. J Subst Abuse Treat 2005 Jun;28(4):305-312 [FREE Full text] [doi: 10.1016/j.jsat.2005.02.004] [Medline: 15925264]
32. Civljak M, Stead LF, Hartmann-Boyce J, Sheikh A, Car J. Internet-based interventions for smoking cessation. Cochrane Database Syst Rev 2013;7:CD007078. [doi: 10.1002/14651858.CD007078.pub4] [Medline: 23839868]
33. D'Amico EJ, Miles JN, Stern SA, Meredith LS. Brief motivational interviewing for teens at risk of substance use consequences: a randomized pilot study in a primary care clinic. J Subst Abuse Treat 2008 Jul;35(1):53-61. [doi: 10.1016/j.jsat.2007.08.008] [Medline: 18037603]
34. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
35. Heatherton TF, Kozlowski LT, Frecker RC, Fagerström K. 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. Bleich S, Havemann-Reinecke U, Kornhuber J. Fagerström-Test für Nikotinabhängigkeit (FTNA). Manual. Göttingen: Beltz Test GmbH; 2002.
37. Annaheim B, Rehm J, Gmel G. How to screen for problematic cannabis use in population surveys: an evaluation of the Cannabis Use Disorders Identification Test (CUDIT) in a Swiss sample of adolescents and young adults. Eur Addict Res 2008;14(4):190-197. [doi: 10.1159/000141643] [Medline: 18583916]

XSL•FO
RenderX

38. Maisto SA, Krenek M, Chung T, Martin CS, Clark D, Cornelius J. Comparison of the concurrent and predictive validity of three measures of readiness to change marijuana use in a clinical sample of adolescents. J Stud Alcohol Drugs 2011 Jul;72(4):592-601 [FREE Full text] [Medline: 21683041]
39. Norberg MM, Mackenzie J, Copeland J. Quantifying cannabis use with the timeline followback approach: a psychometric evaluation. Drug Alcohol Depend 2012 Mar 1;121(3):247-252. [doi: 10.1016/j.drugalcdep.2011.09.007] [Medline: 21955364]
40. Core Team R. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2013. URL: http://www.r-project.org/ [accessed 2014-11-13] [WebCite Cache ID 6U3m4sDap]
41. Blankers M, Koeter MW, Schippers GM. Missing data approaches in eHealth research: simulation study and a tutorial for nonmathematically inclined researchers. J Med Internet Res 2010;12(5):e54 [FREE Full text] [doi: 10.2196/jmir.1448] [Medline: 21169167]
42. StataCorp. Stata Statistical Software: Release. College Station, TX: StataCorp LP; 2011. URL: http://www.stata.com/ [accessed 2014-11-14] [WebCite Cache ID 6U4lsdrPF]
43. IBM SPSS Statistics 2. Chicago, IL: SPSS, Inc; 2011. URL: http://www-01.ibm.com/software/analytics/spss/products/statistics/ [accessed 2014-11-14] [WebCite Cache ID 6U4lWPsPX]
44. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power, version 3.1.3. URL: http://www.gpower.hhu.de/ [accessed 2014-11-14] [WebCite Cache ID 6U4lnxlMa]
45. 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]
46. Tait RJ, Spijkerman R, Riper H. Internet and computer based interventions for cannabis use: a meta-analysis. Drug Alcohol Depend 2013 Dec 1;133(2):295-304. [doi: 10.1016/j.drugalcdep.2013.05.012] [Medline: 23747236]
47. Cohen J. Statistical power analysis for the behavioral sciences. 2nd edition. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
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]
49. Jonas B, Tossmann P, Tensil M, Leuschner F, Strüber E. Effekte einer einmaligen Chat-Intervention auf problematischen Substanzkonsum. SUCHT - Zeitschrift für Wissenschaft und Praxis / Journal of Addiction Research and Practice 2012 Jun 19;58(3):173-182. [doi: 10.1024/0939-5911.a000182]
50. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
51. Woodruff SI, Conway TL, Edwards CC, Elliott SP, Crittenden J. Evaluation of an Internet virtual world chat room for adolescent smoking cessation. Addict Behav 2007 Sep;32(9):1769-1786. [doi: 10.1016/j.addbeh.2006.12.008] [Medline: 17250972]
52. Haug S, Meyer C, John U. Efficacy of an internet program for smoking cessation during and after inpatient rehabilitation treatment: a quasi-randomized controlled trial. Addict Behav 2011 Dec;36(12):1369-1372. [doi: 10.1016/j.addbeh.2011.08.004] [Medline: 21907496]
53. Rooke S, Copeland J, Norberg M, Hine D, McCambridge J. Effectiveness of a self-guided web-based cannabis treatment program: randomized controlled trial. J Med Internet Res 2013;15(2):e26 [FREE Full text] [doi: 10.2196/jmir.2256] [Medline: 23470329]
54. Tossmann HP, Jonas B, Tensil MD, Lang P, Strüber E. A controlled trial of an internet-based intervention program for cannabis users. Cyberpsychol Behav Soc Netw 2011 Nov;14(11):673-679. [doi: 10.1089/cyber.2010.0506] [Medline: 21651419]
55. Lee CM, Neighbors C, Kilmer JR, Larimer ME. A brief, web-based personalized feedback selective intervention for college student marijuana use: a randomized clinical trial. Psychol Addict Behav 2010 Jun;24(2):265-273 [FREE Full text] [doi: 10.1037/a0018859] [Medline: 20565152]
56. McCambridge J, Hunt C, Jenkins RJ, Strang J. Cluster randomised trial of the effectiveness of motivational interviewing for universal prevention. Drug Alcohol Depend 2011 Apr 1;114(2-3):177-184. [doi: 10.1016/j.drugalcdep.2010.07.028] [Medline: 21075562]
57. McCambridge J, Slym RL, Strang J. Randomized controlled trial of motivational interviewing compared with drug information and advice for early intervention among young cannabis users. Addiction 2008 Nov;103(11):1809-1818. [doi: 10.1111/j.1360-0443.2008.02331.x] [Medline: 18778385]

Abbreviations

CUDIT: Cannabis Use Disorder Identification Test
FTND: Fagerstrom Test of Nicotine Dependence
GEE: generalized estimating equations
MI: motivational interviewing (intervention 2)
NF: normative feedback (intervention 1)

XSL•FO
RenderX

PE: psychoeducation (intervention 3)
SE: standard error
t0: baseline assessment
t1: post-intervention assessment
t2: 8-week follow-up assessment

Edited by G Eysenbach; submitted 13.01.14; peer-reviewed by M Blankers, P Hysert; comments to author 26.04.14; revised version received 30.09.14; accepted 25.10.14; published 05.12.14.

Please cite as:
Becker J, Haug S, Sullivan R, Schaub MP
Effectiveness of Different Web-Based Interventions to Prepare Co-Smokers of Cigarettes and Cannabis for Double Cessation: A Three-Arm Randomized Controlled Trial
J Med Internet Res 2014;16(12):e273
URL: http://www.jmir.org/2014/12/e273/
doi:10.2196/jmir.3246
PMID:25486674

©Julia Becker, Severin Haug, Robin Sullivan, Michael Patrick Schaub. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 05.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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

Dose-Response Effects of a Web-Based Physical Activity Program on Body Composition and Metabolic Health in Inactive Older Adults: Additional Analyses of a Randomized Controlled Trial

David P Vroege[1], MD; Carolien A Wijsman[1], MD, PhD; Karen Broekhuizen[2], PhD; Anton JM de Craen[1], PhD; Diana van Heemst[1], PhD; Frans JG van der Ouderaa[3,4], PhD; Willem van Mechelen[5], MD, PhD; P Eline Slagboom[6], PhD; Michael Catt[7], PhD; Rudi GJ Westendorp[3], MD, PhD; Evert ALM Verhagen[5], PhD; Simon P Mooijaart[1,2,4], MD, PhD

[1]Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, Netherlands

[2]Institute for Evidence-Based Medicine in Old Age | IEMO, Leiden, Netherlands

[3]Leyden Academy on Vitality and Ageing, Leiden, Netherlands

[4]Netherlands Consortium for Healthy Aging, Leiden, Netherlands

[5]Department of Public & Occupational Health and EMGO+ Institute, VU University Medical Center, Amsterdam, Netherlands

[6]Department of Medical Statistics, section of Molecular Epidemiology, Leiden University Medical Center, Leiden, Netherlands

[7]Institute for Aging and Health, Newcastle University, Newcastle upon Tyne, United Kingdom

Corresponding Author:
Simon P Mooijaart, MD, PhD
Department of Gerontology and Geriatrics
Leiden University Medical Center
Postzone C7-Q
PO Box 9600
Leiden, 2300 RC
Netherlands
Phone: 31 715266640
Fax: 31 5266881
Email: S.P.Mooijaart@lumc.nl

Abstract

Background: Low physical activity is a major risk factor for several age-related diseases. Recently, we showed in a randomized controlled trial that a 12-week Web-based intervention (Philips DirectLife) to increase physical activity was effective in increasing physical activity levels and metabolic health in an inactive population aged 60-70 years.

Objective: The goal of this paper was to assess how many participants successfully reached the physical activity level as targeted by the intervention and what the effects of the intervention on body composition and metabolic health in these successful individuals were to provide insight in the maximum attainable effect of the intervention.

Methods: Among the 235 participants in a randomized controlled trial of the Actief en Gezond Oud (AGO) study, we assessed the effects of the intervention on metabolic parameters in those who had successfully reached their personalized physical activity target compared with the entire intervention group. Furthermore, we studied the dose-response effect of increase in physical activity on metabolic outcome within the intervention group.

Results: Of the intervention group, 50 of 119 (42.0%) participants successfully reached the physical activity target (corresponding to a 10% increased daily physical activity on average). This group showed markedly higher effects of the intervention compared to the entire intervention group, with greater decreases in body weight (2.74 vs 1.49 kg), waist circumference (3.74 vs 2.33 cm), insulin resistance (HOMA index: 0.23 vs 0.20), and in cholesterol/HDL ratio (0.39 vs 0.20) and Framingham risk score (0.90% vs 0.54%). We found that men compared to women were more likely to be successful. The dose-response analysis showed that there was a significant association between increase in minutes spent in moderate-to-vigorous activity and body weight loss, BMI reduction, waist circumference reduction, HDL cholesterol increasing, and cholesterol/HDL ratio lowering.

Conclusions: Of the intervention group, 42.0% (50/119) reached their daily physical activity end goal, which was associated with a markedly better effect on body composition and metabolic health compared to the effect in the entire intervention group. In this population, men are more likely to be successful in increasing physical activity. Findings demonstrate that improving the

XSL•FO
RenderX

effect of such physical activity interventions requires finding new ways to increase the proportion of the population reaching the targeted goal.

Trial Registration: Dutch Trial Registry: NTR 3045; http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=3045 (Archived by WebCite at http://www.webcitation.org/6KPw52dCc).

(J Med Internet Res 2014;16(12):e265) doi:10.2196/jmir.3643

KEYWORDS

Internet; physical activity; aging; metabolism

Introduction

Insufficient physical activity is pandemic and a major risk factor for several lifestyle- and age-related conditions, including cardiovascular disease, diabetes mellitus, and cognitive decline [1-7]. Intervention studies directed at increasing physical activity in older people have shown to be effective in improving metabolic health in older populations [8,9]. However, most physical activity interventions have used face-to-face communication, making them costly and time-consuming, thus hampering the potential of implementation as preventive programs at a larger scale. There is a need for new and effective intervention strategies that allow for large-scale implementation.

Recently, we performed the Actief en Gezond Oud (AGO) study, a randomized controlled trial of the effect of a 3-month Web-based intervention program targeted at improving physical activity in inactive older adults [10]. The intervention program (Philips DirectLife) consisted of the use of an accelerometer, online feedback, and coaching over the Internet. Results showed that the intervention was effective in increasing objectively measured physical activity and in improving metabolic health in inactive older adults in the total study population. However, this intention-to-treat analysis did not include analyses of treatment success at the level of the individual. Understanding the proportion of the study population that successfully reached the physical activity target and the effects of the intervention in these individuals could contribute to gaining insight in the maximum attainable result of such interventions, allowing future interventions to optimize effectiveness by target specific populations or adjusting the target physical activity level [11,12].

In the present paper, we performed additional analyses on the AGO study data. First, we analyzed what proportion of participants successfully reached the physical activity target of the program. Second, we analyzed the effect on metabolic outcomes of the intervention in those participants who successfully reached the physical activity target. Third, we performed a dose-response analysis of increasing physical activity on metabolic outcomes among all participants in the intervention group.

Methods

Study Design and Participants

All analyses of this paper were performed with data obtained from a previously reported randomized controlled trial into the effects of a 3-month Web-based intervention program targeted at enhancing levels of daily physical activity in inactive older adults: The AGO study [10]. In short, this study recruited participants aged 60 to 70 years from the region of Leiden, The Netherlands, through advertisement in local newspapers and press notification, directing participants motivated to increase physical activity to the study website, where they completed an online questionnaire. Inclusion criteria were (1) age between 60 and 70 years and (2) possession of and knowledge on how to use a personal computer. Exclusion criteria were (1) active lifestyle as assessed by the General Practice Physical Activity Questionnaire (GPPAQ), (2) history of diabetes or use of glucose lowering medication, and (3) physical inability or medical contraindication to increase physical activity level. The presence of an inactive lifestyle was then assessed by a self-report physical activity questionnaire (GPPAQ) [13]. The GPPAQ asks questions about average physical activity over the past week of the participant and categorizes people into 4 levels of physical activity. We excluded participants in the highest level of physical activity, which corresponded to performing more than 3 hours of self-reported exercise and cycling combined weekly. At baseline, visit participants were randomly assigned to the intervention group or a waiting list control. Randomization was performed by a computerized program for intervention versus waiting list control in a ratio of 1:1 with a block size of 12. Stratification was performed by gender. Concealment of treatment allocation was ensured by randomizing at the end of the first study visit after all baseline measurements and instructions at the study center were completed. The study was approved by the medical ethical committee of Leiden University Medical Center, The Netherlands. An independent physician was available for questions regarding study information. This study was registered with the Dutch Trial Registry (NTR 3045).

Intervention

Participants in the intervention group received a commercially available Web-based physical activity program (DirectLife, Philips, Consumer Lifestyle, Amsterdam) directed at increasing daily physical activity. The DirectLife program is based on established health behavior change models [14,15], and takes into account the individual's current daily physical activity level and subsequently provides a personalized goal. Briefly, DirectLife consists of 3 elements: (1) an accelerometer-based physical activity monitor, (2) a personal website, and (3) a personal e-coach, who provides regular updates of the individual's physical activity status by email and who gives advice to increase daily physical activities (Figure 1). By means of these elements, the program aims to increase awareness about one's own physical activity behavior, to give feedback on recent

XSL•FO
RenderX

actual physical activity, and to provide support to make sustainable changes in physical activity behavior.

The activity monitor of DirectLife is based on the Tracmor triaxial accelerometer and has been validated against doubly labeled water for the estimation of total 24-hour energy expenditure [16]. The DirectLife monitor is the consumer version of the Tracmor accelerometer. Intervention group participants received the program, including the accelerometer, directly after randomization at the first study visit. They then received a link by email for registration and access to the Web-based program. Participants were instructed to continuously wear the activity monitor throughout the day to measure daily physical activity. Data were uploaded through an Internet connection to the database of the commercial provider on a regular basis, ranging between daily and once per 14 days. After an initial 8-day "assessment period" starting 1 week after the study visit, in which the current level of daily physical activity was measured, a target was set by the DirectLife program to increase the level of daily physical activity during a 12-week Web-based interactive coaching program. Personalized targets were set by the DirectLife program and were defined as the absolute increase in physical activity compared to the individual's baseline assessment data. For the whole group, this corresponded to a mean increase of approximately 10% in daily physical activity at week 12, increasing at a linear rate per week. All participants were given the option to decrease the personalized goal (within limits: increase of daily physical activity minimally to 5% instead of 10%) or to increase their personalized end goal (dependent on physical activity level of the last week).

Participants were given a target for daily physical activity, which increased weekly, and data from the accelerometer were used for daily feedback. Coaching included general recommendations on physical activities from real-life coaches, who were available for further questions and advice by email correspondence. For every participant, 1 of the e-coaches was available for the DirectLife program during the entire study period. These coaches were actual persons in contact with the participant through the intervention website or through email. The control group was placed on a 3-month waiting list after which they received access to the intervention program at the end of the study. No specific instructions regarding daily physical activity were given.

Figure 1. Screenshot of DirectLife intervention program and accelerometer.



Measurements

Baseline Questionnaire

Enrollment and follow-up took place from November 2011 to August 2012. In preparation of the first visit to the study center, all participants completed a Web-delivered questionnaire on education, smoking status, and medical history, including medication use. Education was categorized as low (primary education and lower vocational education), intermediate (secondary education and intermediate vocational education), or high (high vocational education and university).

XSL•FO
RenderX

Physical Activity Outcome

At baseline and 3-month follow-up, daily physical activity was measured for 7 days following the visit at the study center using a wrist-worn triaxial accelerometer (GeneActiv, Kimbolton, Cambs, UK). Wearing of the GeneActive monitors started on a random weekday depending on which weekday the participant was included for the first visit at the study center, and monitors were returned after 7 days by standard mail. We chose to assess the primary outcome using accelerometers other than the one included in the intervention program to avoid interpretation of the intervention as an outcome. The GeneActive monitors were worn 24 hours per day on the right wrist and ankle. The GeneActiv wrist accelerometer provides a simple summary statistic of total physical activity counts that has been validated for measuring daily physical activity against doubly labeled water [17]. As a derivative outcome, we calculated the minutes per day spent in moderate-to-vigorous intensity physical activity from the wrist accelerometer, which has been validated against indirect calorimetry [18]. Measurement frequency was set at 85.7 Hz and raw acceleration values (in *g*) were recorded continuously on each axis over 7 consecutive days. Further details on data processing can be found elsewhere [10]. Outcome assessment was done by an independent researcher who was blind to study arm allocation.

Other Outcomes

To calculate body mass index (BMI), body height was measured without shoes using a stadiometer. Body weight was assessed at both visits without shoes using a scale. Waist circumference was obtained in a standing position halfway between the anterior superior iliac spine and the lower rib. Hip circumference was measured halfway between the trochanter major and the iliac crest.

Lean body mass and body fat percentage were assessed by bioelectrical impedance analysis (Biostat 1500, Euromedix, Leuven, Belgium). Blood pressure was measured manually twice at each visit using a handheld sphygmomanometer after 5 minutes of lying down. The mean of the 2 consecutive measurements was used. Pulse rate was measured by hand at the wrist after at least 5 minutes of lying down. Grip strength was measured to the nearest kilogram 3 times using a Jamar handheld dynamometer (Sammons Preston, Inc, Bolingbrook, IL, USA) with the dominant hand. The highest value was used for analysis. Framingham risk scores were calculated using NIH criteria [19].

Metabolic Outcomes

Fasting blood samples were drawn from each participant at both visits in the morning. Samples were transferred to the laboratory within 2 hours, divided into single-use aliquots, and frozen at -80 °C. All serum measurements were performed in 1 batch after completion of the entire study with fully automated equipment. Fasting glucose, cholesterol, high-density lipoprotein (HDL) cholesterol, and triglyceride levels were determined using the Modular P2 analyzer (Roche, Almere, the Netherlands), fasting serum insulin using immunoassay by Immulite 2500 (DPC, Los Angeles, CA, USA). Glycated hemoglobin was determined by high performance liquid chromatography (Primus Ultra2, Trinity Biotech Company, Kansas City, MO, USA). C-reactive protein was determined using a high-sensitivity immunoassay (COBAS Integra, Roche, IN, USA). Low-density lipoprotein (LDL) cholesterol was calculated using the Friedewald formula in participants without hypertriglyceridemia [20].

End Point of the DirectLife Intervention

To assess the potential effects of the Web-based intervention, a subgroup was created from the intervention group including participants successful in reaching their individually targeted increase in daily physical activity indicated by the intervention program. In the last 3 weeks of the program, an average physical activity level per week was calculated and was compared to the personalized target of the corresponding week. A participant was defined as successful when the average of at least 2 of the 3 weeks reached their personalized target of the DirectLife program. It was noticed that a substantial number of participants reached the targeted personalized goals at the end of the 12-week program, but there was some variation in the last 3 weeks. Therefore, we defined “successful” as the participants who reached the target in at least 2 of the 3 last weeks of the program.

Statistical Analyses

Baseline differences between the successful participants and the control group and between the entire intervention group and the control group were calculated using a *t* test for continuous data, a Mann-Whitney analysis for skewed data, and a chi-square test for categorical data. Differences between baseline and follow-up within groups were tested using a paired sample Student *t* test of the means. Differences between groups were calculated using linear regression and were adjusted for age and gender. All analyses were performed with SPSS version 20.0 (IBM, Armonk, NY, USA). Statistical significance was set at $P<.05$.

Results

Participant Characteristics

A detailed flow of recruitment and inclusion is outlined elsewhere [10]. In short, a total number of 631 individuals responded to the newspaper advertisement, of which 344 fulfilled the selection criteria. In total, 235 participants were randomized into the study: 119 in the intervention group and 116 in the control group. Of the 235 randomized participants, 226 (96.2%) completed the trial.

Of the 119 participants in the intervention group at baseline, 5 participants were lost to follow-up (4.2% of intervention group). Among the 114 participants who completed the trial, 50 participants (42.0% of intervention group) successfully reached their personalized physical activity target (“successful” participants). Of the 64 who were not defined as successful, 13 did not finish the DirectLife program (10.9% of intervention group) and 51 did not reach the personalized target (42.9% of intervention group).

Table 1 shows the baseline characteristics of the entire intervention group (n=119) and the successful participants (n=50) and both groups were compared with the entire control

XSL•FO
RenderX

group (n=116). Male participants were more likely to successfully reach their personalized target for DirectLife compared to female participants. Of the successful participants, 26% (13/50) were female compared to 39.5% (47/119) in the entire intervention group. In-line with a difference in gender distribution, the average body height of the successful participants was higher compared to the control group. No other significant differences between groups were found.

On average, the personalized goals for males and females were similar. In the total intervention group, 36.1% (43/119) of participants changed their personalized goal: 15.1% (18/119) decreased the goal and 21.0% (25/119) increased the goal. There was no difference between men and women (*P*=.68). In the successful group, 44% (22/50) of the participants changed their personalized goal. There was an overrepresentation of participants increasing their goal (30%, 15/50) and men were more likely to increase their goal (38%, 14/37) compared to women (8%, 1/13; *P*=.038) (data not shown).

Table 1. Baseline characteristic of control group, total intervention group, and successful participants.[a]

Characteristics	Control group	Intervention group			
	Control group (n=116)	Intervention group (n=119)	*P* value for intervention vs control group	Successful participants (n=50)	*P* value for successful participants vs control group
Demographics					
Sex (female), n (%)	49 (42.2)	47 (39.5)	.67	13 (26)	.047
Age (years), mean (SD)	64.9 (2.8)	64.7 (3.0)	.61	64.6 (2.8)	.63
Clinical parameters, mean (SD)					
Height (cm)	172.1 (9.3)	173.6 (9.9)	.25	175.7 (9.6)	.02
Weight (kg)	86.3 (15.8)	87.4 (15.8)	.61	87.6 (15.6)	.64
BMI (kg/m^2)	29.1 (4.7)	28.9 (4.7)	.84	28.2 (3.7)	.25
Waist circumference (cm)	101.4 (12.3)	102.3 (13.1)	.56	102.1 (12.2)	.74
Fat percentage (%)	36.4 (8.1)	36.5 (7.6)	.95	34.5 (6.3)	.11
Cardiovascular disease risk					
Framingham 10-year CVD risk (%), mean (SD)	11.3 (7.5)	11.9 (7.2)	.50	13.3 (7.5)	.10
Physical activity					
5-day moderate-to-vigorous activity (min/day), median (IQR)	14.5 (8.2-32.5)	16.8 (7.8-26.4)	.43	20.0 (9.2-27.5)	.97
Biochemistry					
Fasting venous glucose, mean (SD)	5.7 (0.8)	5.7 (0.7)	.94	5.6 (0.6)	.78
Fasting insulin (mU/L), median (IQR)	10.8 (7.0-15.8)	11.5 (8.1-16.9)	.47	12.4 (8.0-19.8)	.32
HbA_{1c} (%), mean (SD)	5.4 (0.3)	5.4 (0.3)	.44	5.4 (0.2)	.72
HOMA index, median (IQR)	2.6 (1.7-4.3)	2.8 (2.0-4.3)	.48	3.0 (2.0-5.0)	.39
Total cholesterol, mean (SD)	5.8 (1.0)	5.7 (1.1)	.74	5.6 (1.1)	.28
HDL cholesterol, mean (SD)	1.4 (0.4)	1.5 (0.5)	.51	1.4 (0.5)	.93
Triglycerides, median (IQR)	1.4 (1.1-2.0)	1.5 (1.1-2.0)	.65	1.5 (1.1-2.0)	.99
LDL cholesterol, mean (SD)	3.6 (0.9)	3.6 (1.0)	.66	3.4 (1.0)	.25
Total/HDL cholesterol ratio, mean (SD)	4.3 (1.3)	4.2 (1.3)	.65	4.2 (1.3)	.70
C-reactive protein, median (IQR)	1.4 (0.8-4.1)	1.6 (0.8-3.1)	.83	1.7 (0.7-3.8)	.85

[a] Data are presented as medians with interquartile range (IQR) when skewed. *P* values were calculated with *t* test (continuous data), Mann-Whitney (skewed data), or chi-square (categorical data).

XSL•FO
RenderX

Effects on the Successful Participants

Table 2 and Figure 2 show the effects of the intervention at follow-up for the successful participants and the entire intervention group, both compared to the control group. Here we assess the magnitude of the effects in the group of successful participants and the total intervention group compared to the control group. Among the successful participants, time spent in moderate-to-vigorous intensity physical activity was higher (mean 18.8, SE 3.9 min/day) compared to the entire intervention group (mean 11.1, SE 2.1 min/day). The successful participants lost more body weight (mean 2.74, SE 0.40 kg) compared to the entire intervention group (mean 1.49, SE 0.26 kg), an 80% higher decrease. Beneficial effects were also seen for waist circumference with a decrease of mean 3.74 (SE 0.55) cm vs mean 2.33 (SE 0.36) cm in the successful participants vs the entire intervention group (62% higher decrease). In-line with the beneficial changes in body composition, significant improvements were seen in metabolic outcomes in the successful participants compared to the entire intervention group. Beneficial effects were seen for the homeostatic model assessment (HOMA) index, a marker for insulin resistance, with decreases of mean 0.23 (SE 0.06) and mean 0.20 (SE 0.05), respectively (15% higher decrease), and decreases for the cholesterol/HDL ratio of mean 0.39 (SE 0.11) and mean 0.20 (SE 0.07), a 95% higher decrease. For the Framingham risk score, a decrease was seen among the successful participants of 0.90% (SE 0.46) compared to 0.54% (SE 0.33) in the entire intervention group.

In a sensitivity analysis, we repeated all calculations with a different definition of “successful”; namely, only those who reached their personalized target in week 12 (n=40). We found similar results (data not shown).

Table 2. Results for clinical parameters and glucose metabolism of successful participants compared to the total control group.[a]

Characteristics	Control group (n=112)	Total intervention group (n=114)		Successful participants (n=50)		Relative increase[d]
	Mean Δ (SE)	Mean Δ (SE)	*P* [b]	Mean Δ (SE)	*P* [c]	
Clinical parameters						
Weight (kg)	–0.82 (0.21)	–1.49 (0.26)	.05	–2.74 (0.40)	<.001	1.84
BMI (kg/m^2)	–0.29 (0.07)	–0.50 (0.09)	.07	–0.91 (0.13)	<.001	1.82
Waist circumference (cm)	–1.29 (0.34)	–2.33 (0.36)	.04	–3.74 (0.55)	<.001	1.62
Body fat (%)	–0.03 (0.24)	–0.88 (0.28)	.03	–1.33 (0.34)	.001	1.51
Cardiovascular disease risk						
Framingham risk score (%)	–0.01 (0.31)	–0.54 (0.33)	.25	–0.90 (0.46)	.13	1.67
Physical activity						
5-day moderate-to-vigorous activity (min/day)	–0.15 (1.5)	11.1 (2.1)	<.001	18.8 (3.86)	<.001	1.69
Biochemistry						
Fasting venous glucose (mmol/L)	–0.13 (0.04)	–0.20 (0.05)	.32	–0.14 (0.06)	.77	0.70
Ln[e] insulin (mU/L)	–0.04 (0.04)	–0.16 (0.04)	.04	–0.20 (0.06)	.03	1.25
HbA_{1c} (%)	–0.01 (0.01)	–0.05 (0.01)	.048	–0.05 (0.02)	.07	1.00
Ln HOMA index	–0.06 (0.04)	–0.20 (0.05)	.04	–0.23 (0.06)	.05	1.15
Total cholesterol (mmol/L)	–0.18 (0.05)	–0.25 (0.06)	.42	–0.38 (0.09)	.08	1.52
HDL cholesterol (mmol/L)	–0.04 (0.02)	–0.008 (0.02)	.29	0.03 (0.03)	.07	N/A
Ln[e] triglycerides (mmol/L)	–0.06 (0.02)	–0.10 (0.03)	.35	–0.18 (0.04)	.02	1.80
LDL cholesterol (mmol/L)	–0.11 (0.04)	–0.17 (0.04)	.43	–0.28 (0.07)	.08	1.65
Cholesterol/HDL ratio	–0.05 (0.05)	–0.20 (0.07)	.12	–0.39 (0.11)	.007	1.95
Ln[e] C-reactive protein (mg/L)	–0.11 (0.09)	–0.12 (0.08)	.93	–0.24 (0.15)	.50	2.00

[a] *P* values between groups were calculated with linear regression. All *P* values were adjusted for age and sex.

[b] *P* value for total control group vs total intervention group.

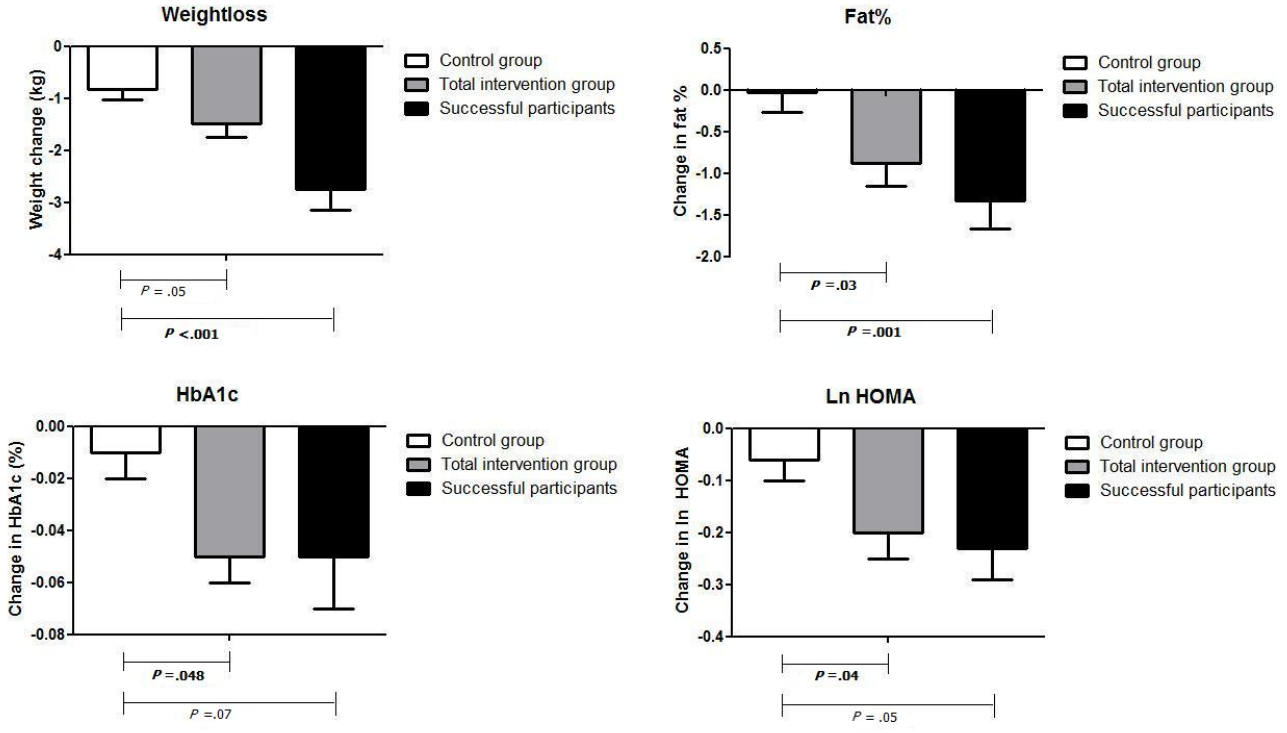
[c] *P* value for total control group vs successful participants.

[d] Relative increase of successful participants compared to total intervention group.

[e] Ln=natural logarithm. Natural logarithm presented when data at baseline skewed.

XSL•FO
RenderX

Figure 2. Effect of the intervention on selected parameters in the different study groups.



Tertiles for Dose-Response Relationship

To explore the effect of physical activity, the entire intervention group was divided into tertiles based on the change in minutes spent in moderate-to-vigorous intensity physical activity. Because of technical errors, data on moderate-to-vigorous intensity physical activity counts were not available for 11 of 119 intervention group participants (9.2%), resulting in 3 tertiles with 36 participants. The lowest tertile showed, on average, a decrease of 6.75 (SD 8.20) minutes spent in moderate-to-vigorous intensity physical activity. The middle tertile showed on average an increase of 5.91 (SD 3.36) minutes and the highest tertile showed on average an increase of 34.37 (SD 21.52) minutes spent in moderate-to-vigorous intensity physical activity.

Dose-Response Relationship of Physical Activity on Metabolic Outcomes

Finally, we assessed the association of the increase in physical activity levels with metabolic outcomes in a dose-response relationship. Table 3 shows the tertiles that were made based on increase in minutes spent in moderate-to-vigorous intensity physical activity as objectively measured with the wrist-worn accelerometer. At baseline, no differences were found between tertiles with regard to physical activity levels. When change in minutes spent in moderate-to-vigorous activity increased, more participants were defined successful (*P* for trend =.001). Furthermore, it was shown that there was a significant association of increasing physical activity with decreasing body weight (highest tertile: mean –2.85, SE 0.51 kg; lowest tertile: mean –0.93, SE 0.35 kg; *P* for trend=.004), decreasing BMI (highest tertile: mean –0.88, SE 0.17 kg/m^2; lowest tertile: mean –0.34 SE 0.12 kg/m^2; *P* for trend=.02), and reduction in waist circumference (highest tertile: mean –3.69, SE 0.72 cm; lowest tertile: mean –1.58, SE 0.52 cm; *P* for trend=.03). Furthermore, it was shown that with increasing physical activity, there was a higher improvement in Framingham risk score (highest tertile: mean –1.28%, SE 0.76%; lowest tertile: mean 0.43%, SE 0.46%; *P* for trend=.045). Also for metabolic outcomes, a dose-response was seen for (levels of) HDL cholesterol, cholesterol/HDL ratio, and triglycerides, but not for fasting glucose or HbA_{1c}.

XSL•FO
RenderX

Table 3. Dose-response relationship of Δ minutes spent in moderate-to-vigorous physical activity with endpoints.[a]

Characteristics	Intervention group			*P* for trend
	Low (n=36)	Middle (n=36)	High (n=36)	
Baseline moderate-to-vigorous activity				
Mean (SD)	22.5 (17.9)	17.3 (16.5)	19.0 (14.9)	.38
Range	2.4 to 87.2	0.8 to 92.0	2.4 to 62.4	
Δ minutes moderate-to-vigorous activity				
Mean (SD)	–6.75 (8.20)	5.91 (3.56)	34.27 (21.52)	
Range	–25.8 to 1.60	1.70 to 14.2	15.8 to 117.6	
Successful participants, n	11	12	26	.001
Clinical parameters (mean Δ, SE)				
Weight (kg)	–0.93 (0.35)	–0.64 (0.44)	–2.85 (0.51)	.004
BMI (kg/m^2)	–0.34 (0.12)	–0.23 (0.15)	–0.88 (0.17)	.02
Waist circumference (cm)	–1.58 (0.52)	–1.92 (0.67)	–3.69 (0.72)	.03
Fat percentage	–0.46 (0.48)	–0.44 (0.35)	–0.94 (0.43)	.43
Cardiovascular disease risk (mean Δ, SE)				
Framingham risk score (%)	0.43 (0.46)	–0.78 (0.52)	–1.28 (0.76)	.045
Biochemistry (mean Δ, SE)				
Glucose (mmol/L)	–0.23 (0.10)	–0.15 (0.09)	–0.21 (0.07)	.94
Ln insulin (mU/L)	–0.14 (0.07)	–0.06 (0.08)	–0.26 (0.08)	.20
HbA_{1c} (%)	–0.04 (0.02)	–0.03 (0.02)	–0.07 (0.02)	.20
Ln HOMA index	–0.18 (0.08)	–0.09 (0.08)	–0.30 (0.08)	.26
Total cholesterol (mmol/L)	–0.29 (0.08)	–0.10 (0.10)	–0.40 (0.11)	.63
HDL cholesterol (mmol/L)	–0.10 (0.04)	0.01 (0.04)	0.05 (0.03)	.007
Ln triglycerides (mmol/L)	0.003 (0.05)	–0.06 (0.05)	–0.22 (0.07)	.02
LDL cholesterol (mmol/L)	–0.19 (0.06)	–0.09 (0.08)	–0.27 (0.09)	.61
Cholesterol/HDL ratio	0.006 (0.08)	–0.15 (0.11)	–0.48 (0.17)	.02
Ln C-reactive protein (mg/L)	–0.11 (0.10)	–0.17 (0.13)	–0.23 (0.17)	.66

[a]*P* for trend was calculated with linear regression. *P* for trend was adjusted for sex and age, except for Framingham risk score for which age and sex are integrated.

Discussion

Principal Results and Comparison to Prior Work

The findings of the present study are fourfold. First, 42.0% (50/119) of the intervention group had reached their personalized target for daily physical activity. Second, we found almost doubled effects in body weight, BMI, and cholesterol/HDL ratio in those who successfully had reached their personalized daily physical activity target compared to the entire intervention group. Third, we found that men were more successful on reaching the personalized targets than women. Fourth, in the entire intervention group we found that metabolic outcome improved with increasing minutes spent on physical activity.

So far, studies have reported different results on improving daily physical activity in different age groups through a Web-based intervention. Limited studies report on elderly participants from the general population with regard to BMI, sedentary lifestyle, and comorbidities. Compared to a waiting list control group some studies reported an increase in moderate and vigorous intensity physical activity [21] or moderate intensity physical activity and walking [22], whereas other studies reported no significant differences in physical activity [23,24]. All these studies used self-report questionnaires for reporting on physical activity instead of objectively measured physical activity, making it hazardous to assess who had actually reached the targeted increase in physical activity. We report here that 42% of the entire intervention group had reached their personalized daily physical activity target when measured objectively. To our knowledge, data on participants successfully increasing objectively measured physical activity as targeted are very limited. The main reason for this is that very few of the studied interventions mention an objectively measured target increase. A randomized controlled trial performed among 1071 participants (mean age 53 years; 57% of participants with a BMI ≥25) studied the effect on nutrition and physical activity

XSL•FO
RenderX

of a 12-week Internet program called Guide to Health (GTH only) that focused on nutrition and physical activity, the Guide to Health program plus a series of group-based support sessions (GTH+), and a waiting list control group. Physical activity was measured using a pedometer, counting steps/day with a target for all groups to increase physical activity with 2142 steps/day at posttreatment compared to baseline. In the GTH+ group, 41.7% increased the step count as targeted compared to the control group (24.4%, *P*=.07), showing that the GTH+ group tended to be more likely to reach the step goal. The GTH only group showed that 35.8% successfully increased their step count [25]. Although the physical activity in this study was objectively measured in an alternate way, our results show a comparable success rate.

Our study showed many beneficial effects on body composition and metabolic outcomes, with a significant average body weight loss of 2.74 kg in 3 months for the successful participants. Furthermore, our trial showed beneficial effects on body composition and metabolic outcomes also in the control group. The latter finding indicates that we have selected a motivated study population who, while on the waiting list for the intervention, may have adopted other strategies to improve their level of daily physical activity. Furthermore, the finding stresses the importance of a well-chosen control arm in clinical trials. Trials of other Web-based physical activity interventions on body composition showed various results [25-27]. For example, the aforementioned Guide to Health study comparing 2 Web-based interventions directed at healthy nutrition and physical activity reported small effects in randomly assigned individuals of -0.10 kg or -0.25 kg on body weight, none of them significant at long-term follow-up [25]. Other studies reporting the results of non-Web-based interventions directed at improving physical health in sedentary obese elderly people showed a mean body weight loss of 1.8 kg [28] or a mean body weight loss -3.6% [29]. In view of these results, our study showed a large effect on body weight. However, for studies primarily directed at body weight loss (including behavioral components other than physical activity), larger effects were seen compared to our study and most other studies directed at improving daily physical activity [30].

In the present study, we showed a general 1.2-fold to 2-fold larger effect in the successful participants compared to the entire intervention in most parameters. In the dose-response analysis, we observed that with tertiles of increasing objectively measured physical activity, there was an increase in beneficial effects of the intervention on body composition and metabolic outcomes. Of note, those with the lowest baseline physical activity level had the highest increase in physical activity level. The fact that there is a linear relationship between the increase in objectively measured physical activity and beneficial changes in body composition and metabolic outcomes indicates that there is also a beneficial effect of the intervention in participants who did not reach the targeted physical activity goal. An interpretation may be that intervention programs should focus on how to increase the compliance of participants to the program and to set feasible physical activity goals leading to higher physical activity levels in general for all participants.

Our analyses showed men were more likely to be successful compared to women; when successful, men were more likely to increase the personalized target compared to women. These are interesting findings. It is known that studies on the effects of physical activity interventions were mainly conducted in females, emphasizing the need for data about physical activity interventions targeted to men [31]. A possible explanation for our finding could be our recruitment strategy has unintentionally selected for more motivated men compared to the recruited women. This is also reflected in the fact that we recruited more males than females overall. Because we found the successful group to have a higher chance of increasing the target, this may reflect an overrepresentation of highly motivated participants in the successful group. In this group, men were more likely to increase their goal, suggesting that men were more motivated then women.

Additionally, it is likely that in this age group, men are more likely than women to already have adopted more digital and electronic gadgets to their lifestyle. For instance, in this age group it is likely that the men had more experience with working with computers than women because men in this age category were more likely to have a job with computers compared to women, especially because it was a highly educated group. Alternatively, using digital solutions and personalized targets may be more effective in men compared to women in contrast to working out in groups, for instance. This may be the case in our group of highly educated, physically inactive men. Future research is needed to confirm these speculations in this subgroup.

Although it was shown that men were more likely to successfully reach the targeted physical activity level compared to women, it was not the primary goal of the present paper to analyze the determinants of parameters that determine which factors predict which participants were successful in reaching their individualized physical activity target. Such a determinant analysis will be performed in forthcoming studies and will address different research questions with the ultimate aim to better target different populations.

Our analyses show a dose-response relationship between increasing physical activity on metabolic outcome. One consequence of this may be that reaching personalized targets is a useful way to stimulate individuals to adopt a lifestyle with a higher level of physical activity, but that the actual level of physical activity increase is less important. A flexible goal (as used in the present program) may, therefore, be very useful to keep participants motivated: even if the increase in physical activity is less than the initial target, there are still metabolic benefits

In the present paper, we stratified our main analyses into those who were successful versus those who were not and performed a dose-response analysis. We found that for those who were successful, the health benefits were larger than in the total intervention group. Furthermore, we report on gender as a determinant of success; men were more likely to increase their personalized goal, suggesting that men were more motivated. Although the intention-to-treat analysis of the total intervention group shows that the intervention is effective, our stratified

XSL•FO
RenderX

analysis shows that there is a subgroup in whom the effects are much bigger and what characterizes these participants. This identifies whom to target with the current intervention, what other groups need an adjusted targeting, and what the effects of the intervention may be when applied optimally in the optimal population. All these lessons may help in the design and adjustment of the present and comparable interventions, especially in the targeted aging population

Limitations

A drawback of our study is that we selected highly motivated participants who were able to use the Internet, leading to a population with a high education level. This hampers generalizability. Furthermore, men were more likely to reach the personalized physical activity goals successfully making the absolute results of the successful participants more difficult to interpret. Furthermore, we did not measure the food intake, which can be of great importance when studying the effect of the intervention. Food intake may have a direct effect on energy balance and several of the metabolic outcomes. Finally, in the present analyses we did not aim to investigate the determinants of who was going to be successful. Such analyses would, for instance, involve a determinant and cluster analyses and is the topic of ongoing analyses. A strength of this study is that we objectively measured physical activity. Furthermore, although participants were generally overweight, our study population consisted of volunteers in which comorbidities were present, increasing the generalizability toward the general elderly population. Furthermore, this study shows that men were more likely to successfully reach their personalized goals, which is a promising result for future interventions directed at improving physical activity. Furthermore, this study is unique in analyzing the dose-response relationship of physical activity within the intervention group leading to new insights for intervention programs. Finally, because the intervention is a Web-based intervention, it is likely to have a better cost benefit compared to face-to-face physical activity interventions.

Conclusions

In conclusion, 42.0% (50/119) of the intervention group reached the end goal for daily physical activity, which was associated with a markedly better effect on metabolic outcomes compared to the effect in the entire intervention group. In this population, men were more successful at reaching the personalized physical activity targets. Findings demonstrate the large potential of Web-based interventions for improving health in the aging population by increasing physical activity, with possibilities for future improvements in increasing the proportion of the population reaching the targeted physical activity goal.

Acknowledgments

We thank Marjan van der Elst, Robert du Puy, Leonieke ten Brinke, Marja Kersbergen, and Margo van Schie for their valuable contributions. The Actief en Gezond Oud study was financially supported by Philips Consumer Lifestyle, and the Netherlands Genomics Initiative/Netherlands Organization for scientific research (NGI/NWO; 05040202 and 050-060-810). The funders had no role in the design and performance of the study, nor in the analyses or interpretation of the data or in the drafting of the manuscript. DPV and SPM take full responsibility for the integrity of the data.

Authors' Contributions

The authors contributed in the following ways: CAW, RGJW, EV, AJMdC, WvM, FvdO, and SPM designed the study. RGJW, FvdO, and SPM acquired funding. MC calculated accelerometer data. CAW and SPM coordinated participant recruitment, inclusion, and retention. DPV, CAW, AJMdC, and SPM performed statistical analyses. DPV, CAW, RGJW, EV, MC, PS, AJMdC, WvM, KB, DvH, FvdO, and SPM were involved in interpretation of the results and drafting of the manuscript. All authors reviewed the final version of the manuscript and agreed to its submission.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.2 [32].

[PDF File (Adobe PDF File), 993KB - jmir_v16i12e265_app1.pdf]

References

1. Shortreed SM, Peeters A, Forbes AB. Estimating the effect of long-term physical activity on cardiovascular disease and mortality: evidence from the Framingham Heart Study. Heart 2013 May;99(9):649-654. [doi: 10.1136/heartjnl-2012-303461] [Medline: 23474622]
2. Tanasescu M, Leitzmann MF, Rimm EB, Hu FB. Physical activity in relation to cardiovascular disease and total mortality among men with type 2 diabetes. Circulation 2003 May 20;107(19):2435-2439 [FREE Full text] [doi: 10.1161/01.CIR.0000066906.11109.1F] [Medline: 12719277]
3. Manson JE, Nathan DM, Krolewski AS, Stampfer MJ, Willett WC, Hennekens CH. A prospective study of exercise and incidence of diabetes among US male physicians. JAMA 1992 Jul 1;268(1):63-67. [Medline: 1608115]

XSL•FO
RenderX

4. Buchman AS, Boyle PA, Yu L, Shah RC, Wilson RS, Bennett DA. Total daily physical activity and the risk of AD and cognitive decline in older adults. Neurology 2012 Apr 24;78(17):1323-1329 [FREE Full text] [doi: 10.1212/WNL.0b013e3182535d35] [Medline: 22517108]
5. Yaffe K, Barnes D, Nevitt M, Lui LY, Covinsky K. A prospective study of physical activity and cognitive decline in elderly women: women who walk. Arch Intern Med 2001 Jul 23;161(14):1703-1708. [Medline: 11485502]
6. Matthews CE, George SM, Moore SC, Bowles HR, Blair A, Park Y, et al. Amount of time spent in sedentary behaviors and cause-specific mortality in US adults. Am J Clin Nutr 2012 Feb;95(2):437-445 [FREE Full text] [doi: 10.3945/ajcn.111.019620] [Medline: 22218159]
7. Shaw BA, Agahi N. A prospective cohort study of health behavior profiles after age 50 and mortality risk. BMC Public Health 2012;12:803 [FREE Full text] [doi: 10.1186/1471-2458-12-803] [Medline: 22989155]
8. Dubbert PM, Cooper KM, Kirchner KA, Meydrech EF, Bilbrew D. Effects of nurse counseling on walking for exercise in elderly primary care patients. J Gerontol A Biol Sci Med Sci 2002 Nov;57(11):M733-M740. [Medline: 12403802]
9. Fujita K, Nagatomi R, Hozawa A, Ohkubo T, Sato K, Anzai Y, et al. Effects of exercise training on physical activity in older people: a randomized controlled trial. J Epidemiol 2003 Mar;13(2):120-126 [FREE Full text] [Medline: 12675121]
10. 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;15(11):e233 [FREE Full text] [doi: 10.2196/jmir.2843] [Medline: 24195965]
11. Steckler A, Linnan L. Process Evaluation for Public Health Interventions and Research. San Francisco, CA: Jossey-Bass; 2002.
12. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am J Public Health 1999 Sep;89(9):1322-1327. [Medline: 10474547]
13. Physical Activity Policy, Health Improvement Directorate. The General Practice Physical Activity Questionnaire (GPPAQ): a screening tool to assess adult physical activity level, within primary care. 2013. URL: http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_063812 [accessed 2013-07-23] [WebCite Cache ID 6IKT5JweI]
14. Prochaska JO, Velicer WF, Rossi JS, Goldstein MG, Marcus BH, Rakowski W, et al. Stages of change and decisional balance for 12 problem behaviors. Health Psychol 1994 Jan;13(1):39-46. [Medline: 8168470]
15. De Vries H, Mudde AN, Dijkstra A, Willemsen MC. Differential beliefs, perceived social influences, and self-efficacy expectations among smokers in various motivational phases. Prev Med 1998;27(5 Pt 1):681-689. [doi: 10.1006/pmed.1998.0344] [Medline: 9808799]
16. Bonomi AG, Plasqui G, Goris AH, Westerterp KR. Estimation of free-living energy expenditure using a novel activity monitor designed to minimize obtrusiveness. Obesity (Silver Spring) 2010 Sep;18(9):1845-1851. [doi: 10.1038/oby.2010.34] [Medline: 20186133]
17. van Hees VT, Renström F, Wright A, Gradmark A, Catt M, Chen KY, et al. Estimation of daily energy expenditure in pregnant and non-pregnant women using a wrist-worn tri-axial accelerometer. PLoS One 2011;6(7):e22922 [FREE Full text] [doi: 10.1371/journal.pone.0022922] [Medline: 21829556]
18. Esliger DW, Rowlands AV, Hurst TL, Catt M, Murray P, Eston RG. Validation of the GENEA Accelerometer. Med Sci Sports Exerc 2011 Jun;43(6):1085-1093. [doi: 10.1249/MSS.0b013e31820513be] [Medline: 21088628]
19. National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation,Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. Circulation 2002 Dec 17;106(25):3143-3421 [FREE Full text] [Medline: 12485966]
20. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. Clin Chem 1972 Jun;18(6):499-502 [FREE Full text] [Medline: 4337382]
21. Plotnikoff RC, McCargar LJ, Wilson PM, Loucaides CA. Efficacy of an E-mail intervention for the promotion of physical activity and nutrition behavior in the workplace context. Am J Health Promot 2005;19(6):422-429. [Medline: 16022206]
22. Napolitano MA, Fotheringham M, Tate D, Sciamanna C, Leslie E, Owen N, et al. Evaluation of an internet-based physical activity intervention: a preliminary investigation. Ann Behav Med 2003;25(2):92-99. [Medline: 12704010]
23. Kosma M, Cardinal BJ, McCubbin JA. A pilot study of a web-based physical activity motivational program for adults with physical disabilities. Disabil Rehabil 2005 Dec 15;27(23):1435-1442. [doi: 10.1080/09638280500242713] [Medline: 16418058]
24. Slootmaker SM, Chinapaw MJ, Schuit AJ, Seidell JC, Van Mechelen W. Feasibility and effectiveness of online physical activity advice based on a personal activity monitor: randomized controlled trial. J Med Internet Res 2009;11(3):e27 [FREE Full text] [doi: 10.2196/jmir.1139] [Medline: 19674956]
25. Winett RA, Anderson ES, Wojcik JR, Winett SG, Bowden T. Guide to health: nutrition and physical activity outcomes of a group-randomized trial of an Internet-based intervention in churches. Ann Behav Med 2007 Jun;33(3):251-261. [doi: 10.1080/08836610701358045] [Medline: 17600452]

XSL•FO
RenderX

26. Reijonsaari K, Vehtari A, Kahilakoski OP, van Mechelen W, Aro T, Taimela S. The effectiveness of physical activity monitoring and distance counseling in an occupational setting - results from a randomized controlled trial (CoAct). BMC Public Health 2012;12:344 [FREE Full text] [doi: 10.1186/1471-2458-12-344] [Medline: 22578104]
27. Ware LJ, Hurling R, Bataveljic O, Fairley BW, Hurst TL, Murray P, et al. Rates and determinants of uptake and use of an internet physical activity and weight management program in office and manufacturing work sites in England: cohort study. J Med Internet Res 2008;10(4):e56 [FREE Full text] [doi: 10.2196/jmir.1108] [Medline: 19117828]
28. Kallings LV, Sierra Johnson J, Fisher RM, Faire U, Ståhle A, Hemmingsson E, et al. Beneficial effects of individualized physical activity on prescription on body composition and cardiometabolic risk factors: results from a randomized controlled trial. Eur J Cardiovasc Prev Rehabil 2009 Feb;16(1):80-84. [doi: 10.1097/HJR.0b013e32831e953a] [Medline: 19237997]
29. Avila JJ, Gutierres JA, Sheehy ME, Lofgren IE, Delmonico MJ. Effect of moderate intensity resistance training during weight loss on body composition and physical performance in overweight older adults. Eur J Appl Physiol 2010 Jun;109(3):517-525. [doi: 10.1007/s00421-010-1387-9] [Medline: 20169360]
30. Weinheimer EM, Sands LP, Campbell WW. A systematic review of the separate and combined effects of energy restriction and exercise on fat-free mass in middle-aged and older adults: implications for sarcopenic obesity. Nutr Rev 2010 Jul;68(7):375-388. [doi: 10.1111/j.1753-4887.2010.00298.x] [Medline: 20591106]
31. Waters LA, Galichet B, Owen N, Eakin E. Who participates in physical activity intervention trials? J Phys Act Health 2011 Jan;8(1):85-103. [Medline: 21297189]
32. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

AGO: Actief en Gezond Oud
BMI: body mass index
CVD: cardiovascular disease
GPPAQ: General Practice Physical Activity Questionnaire
HbA1c: glycated hemoglobin
HOMA: homeostatic model assessment
HDL: high-density lipoprotein
LDL: low-density lipoprotein

Edited by G Eysenbach; submitted 26.06.14; peer-reviewed by A Booth, C Short; comments to author 24.07.14; revised version received 18.09.14; accepted 22.09.14; published 04.12.14.

Please cite as:
Vroege DP, Wijsman CA, Broekhuizen K, de Craen AJM, van Heemst D, van der Ouderaa FJG, van Mechelen W, Slagboom PE, Catt M, Westendorp RGJ, Verhagen EALM, Mooijaart SP
Dose-Response Effects of a Web-Based Physical Activity Program on Body Composition and Metabolic Health in Inactive Older Adults: Additional Analyses of a Randomized Controlled Trial
J Med Internet Res 2014;16(12):e265
URL: http://www.jmir.org/2014/12/e265/
doi:10.2196/jmir.3643
PMID:25486673

©David P Vroege, Carolien A Wijsman, Karen Broekhuizen, Anton JM de Craen, Diana van Heemst, Frans JG van der Ouderaa, Willem van Mechelen, P Eline Slagboom, Michael Catt, Rudi GJ Westendorp, Evert ALM Verhagen, Simon P Mooijaart. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 04.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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

Reducing Alcohol Use During Pregnancy Via Health Counseling by Midwives and Internet-Based Computer-Tailored Feedback: A Cluster Randomized Trial

Nickie Y van der Wulp[1,2], MSc; Ciska Hoving[2,3], PhD; Kim Eijmael[2,3], MSc; Math JJM Candel[2,4], PhD; Wim van Dalen[1], MSc; Hein De Vries[2,3], PhD

[1]Dutch Institute for Alcohol Policy STAP, Utrecht, Netherlands

[2]Care and Public Health Research Institute, Maastricht University, Maastricht, Netherlands

[3]Department of Health Promotion, Maastricht University, Maastricht, Netherlands

[4]Department of Methodology and Statistics, Maastricht University, Maastricht, Netherlands

Corresponding Author:
Nickie Y van der Wulp, MSc
Dutch Institute for Alcohol Policy STAP
P.O. Box 9769
Utrecht, 3506 GT
Netherlands
Phone: 31 306565041
Fax: 31 306565043
Email: nvanderwulp@stap.nl

Abstract

Background: Effective interventions are needed to reduce neurobehavioral impairments in children due to maternal alcohol use during pregnancy. Currently, health-counseling interventions have shown inconsistent results to reduce prenatal alcohol use. Thus, more research using health counseling is needed to gain more knowledge about the effectiveness of this type of intervention on reducing alcohol use during pregnancy. An alternative and promising strategy is computer tailoring. However, to date, no study has shown the effectiveness of this intervention mode.

Objective: The aim was to test the effectiveness of health counseling and computer tailoring on stopping and reducing maternal alcohol use during pregnancy in a Dutch sample of pregnant women using alcohol.

Methods: A total of 60 Dutch midwifery practices, randomly assigned to 1 of 3 conditions, recruited 135 health counseling, 116 computer tailoring, and 142 usual care respondents from February to September 2011. Health-counseling respondents received counseling from their midwife according to a health-counseling protocol, which consisted of 7 steps addressed in 3 feedback sessions. Computer-tailoring respondents received usual care from their midwife and 3 computer-tailored feedback letters via the Internet. Usual care respondents received routine alcohol care from their midwife. After 3 and 6 months, we assessed the effect of the interventions on alcohol use.

Results: Multilevel multiple logistic regression analyses showed that computer-tailoring respondents stopped using alcohol more often compared to usual care respondents 6 months after baseline (53/68, 78% vs 51/93, 55%; P=.04). Multilevel multiple linear regression analyses showed that computer-tailoring respondents (mean 0.35, SD 0.31 units per week) with average (P=.007) or lower (P<.001) alcohol use before pregnancy or with average (P=.03) or lower (P=.002) social support more strongly reduced their alcohol use 6 months after baseline compared to usual care respondents (mean 0.48, SD 0.54 units per week). Six months after baseline, 72% (62/86) of the health-counseling respondents had stopped using alcohol. This 17% difference with the usual care group was not significant.

Conclusions: This is the first study showing that computer tailoring can be effective to reduce alcohol use during pregnancy; health counseling did not effectively reduce alcohol use. Future researchers developing a health-counseling intervention to reduce alcohol use during pregnancy are recommended to invest more in recruitment of pregnant women and implementation by health care providers. Because pregnant women are reluctant to disclose their alcohol use to health professionals and computer tailoring preserves a person's anonymity, this effective computer-tailoring intervention is recommended as an attractive intervention for pregnant women using alcohol.

XSL•FO
RenderX

Trial Registration: Dutch Trial Register NTR 2058; http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=2058 (Archived by WebCite at http://www.webcitation.org/6NpT1oHol)

(J Med Internet Res 2014;16(12):e274) doi:10.2196/jmir.3493

KEYWORDS

alcohol drinking; pregnancy; counseling; telemedicine; midwifery

Introduction

Alcohol use in pregnancy is a leading preventable cause of intellectual disability in children [1]. Due to accumulating evidence that even low levels of prenatal alcohol exposure can cause adverse neurobehavioral effects in children [2], many Western countries, such as the United States, Australia, and the Netherlands, officially recommend that pregnant women completely abstain from alcohol [3-5]. Nevertheless, more than 20% of pregnant women worldwide consume alcohol [6], with estimations accumulating to 35% to 50% in the Netherlands [4]. Effective interventions are needed to reduce the number of pregnant women who endanger the health of their fetuses by using alcohol in pregnancy.

Various interventions to reduce prenatal alcohol use have been described in reviews [7-9]; however, only 5 studies used a randomized controlled trial to test intervention effectiveness (ie, [10-14]). All these interventions applied health counseling; pregnant women were screened for alcohol use and participated in motivational interviews conducted by health professionals (eg, [15]). These studies suggest that health-counseling interventions may result in increased abstinence and a reduction in prenatal alcohol consumption. However, because of the inconsistency of the results, the paucity of studies, the relatively low number of total respondents, the high risk of bias of the studies due to lack of information on allocation concealment, and the complexities of interventions, many uncertainties remain about the most optimal conditions of these interventions [9].

An alternative and promising strategy is computer tailoring, an intervention in which advice is not delivered face-to-face, but via a computer [16]. The content of this advice is based on the answers of respondents to questions and is generated by a computer program. Consequently, the feedback is adapted to the specific characteristics of a particular individual, yielding the potential to provide messages highly tailored to the individuals' situation [17]. Computer-tailored messages have been shown to attract and keep an individual's attention [16,18] more than generic advice, resulting in a more thorough processing of information [19]. Computer tailoring has proven to be effective in changing different health-related behaviors [20,21], such as smoking [22], vegetable and fruit intake [23], and alcohol use [24,25]. To our knowledge, only 1 computer-tailored intervention for alcohol use in pregnancy has been developed. Tzilos and colleagues [26] found that users liked the program and appreciated the ease of use. Nevertheless, they did not find any significant difference in the reduction of alcohol use compared to an assessment-only condition, perhaps because their 1-month follow-up was simply too soon to find beneficial effects of the computer tailoring or because their phone-based follow-up led to a social desirability bias concealing real decreases in drinking. Thus, it has not been shown that computer tailoring can be effective in reducing prenatal alcohol use.

The goal of this study was to test the effectiveness of 2 different brief interventions to reduce prenatal alcohol use, a health counseling and a computer-tailored intervention, in comparison with usual care. In agreement with several national recommendations [3-5], our primary focus for the development of the interventions was that pregnant women who used alcohol in the beginning of their pregnancy stopped their alcohol use after having received an intervention. Thus, our first hypothesis was that women receiving health counseling or computer tailoring were more likely to stop using alcohol in pregnancy compared to women receiving usual care. However, for the pregnant women unwilling or unable to completely stop their alcohol use, we aimed at reducing their alcohol use because research has shown that the risk and severity of the effects of prenatal alcohol use are dose-related [27]. Consequently, our second hypothesis is that when women continued their alcohol use, those receiving health counseling or computer tailoring were more likely to reduce their alcohol use compared to those receiving usual care.

Methods

Ethical Approval and Registration

The study was approved by the Medical Ethics Committee of Maastricht University and the University Hospital Maastricht (MEC 09-3-070) and is registered with the Dutch Trial Register (NTR2058).

Sample

A sample size analysis, with power=.80, alpha=.05, intraclass correlation coefficient (ICC) of 0.01 (reported in a previous study as the median ICC for cluster-based studies in primary care [28]), an estimated quit rate of 40% in each of the experimental conditions versus 20% in the control condition, and the estimated inclusion of 30 midwifery practices, revealed that 9 respondents per practice were needed. Estimating 10% attrition over the trial period, we aimed to include 300 respondents at baseline. The estimated quit rate and attrition were based on a previous Dutch study on smoking cessation during pregnancy [29].

Eligibility criteria were ability to understand Dutch, aged 18 years or older, pregnant for a maximum of 12 weeks (because respondents received follow-up questionnaires until 6 months after baseline), and having drunk alcohol since knowing to be pregnant.

XSL•FO
RenderX

Procedures

Respondents were recruited from February to September 2011. Recruitment letters were sent to all midwifery practices in the Netherlands (N=540). Participating practices were randomly assigned to 1 of the 3 conditions (health counseling, computer tailoring, or usual care) by a computer software randomization device to avoid contamination. The practices informed their clients about the study by email or phone. When pregnant women agreed to participate, they were asked to visit the study website before their initial consultation. They could do this where and whenever they had access to the Internet.

During the recruitment period, it appeared that the inclusion of 9 respondents per practice would be too time-consuming. We decided to enroll 60 midwifery practices in total, expecting to recruit 4-5 respondents from each practice.

The study website included the baseline questionnaire (T0). Respondents could choose their own username and password and had to report their email address when signing up for the study. This way we could easily remove respondents with multiple identities from further analyses. Before providing informed consent, pregnant women were informed about the 3 study conditions and received information about the objectives of the study, the randomization procedure, and the incentive of a €10 voucher when respondents completed all questionnaires and institutional affiliations ("This research is conducted by the Dutch Institute for Alcohol Policy [STAP] and Maastricht University"). After providing online informed consent, eligible women gained access to the baseline questionnaire. Blinding of respondents was not possible because they had to take notice of whether they did or did not receive additional counseling from their midwife (after the baseline questionnaire) or tailored feedback via the computer (during the baseline questionnaire).

At both 3 and 6 months after the baseline questionnaire, all participants received an invitation by email (followed by 2 reminders after 2 and 4 weeks) for the first follow-up questionnaire (T1 and T2, respectively). Nonrespondents after 2 reminders were contacted by telephone to collect their data.

Interventions

Overview

The health counseling and computer-tailoring interventions were both based on the I-Change model [30], a theoretical model incorporating concepts from several social cognitive models, such as the Transtheoretical model [31] and the Theory of Planned Behavior [32]. The I-Change model distinguishes 3 phases of health behavior change (awareness, motivation, and action) and has been used successfully for developing various health promoting interventions, such as prenatal smoking cessation [29], smoking cessation [22,33,34], and increasing vegetable and fruit intake and physical activity [21].

Health Counseling

Midwives in the health-counseling condition received a brief manual explaining the health-counseling protocol and an intervention card with questions for the clients. On this intervention card, midwives could record the dates of the health counseling sessions and the clients' answers to the midwife's questions. Midwives received 3 hours of training on how to provide the health counseling. This training was given either at the research institute of the first author or at the practice of the participating midwife. The materials and training were based on earlier work on tobacco and pregnancy [29].

The health-counseling protocol consisted of 7 steps which were addressed in 3 feedback sessions. Feedback session 1, approximately 2 weeks after baseline assessment, consisted of 5 steps taking approximately 10 minutes of the initial consultation (Feedback 1-health counseling). In step 1, the midwife assessed the amount and frequency of alcohol use of the pregnant woman before and during pregnancy, of her partner during pregnancy, and the pregnant woman's motivation to stop drinking alcohol. In step 2, women strongly motivated to stop alcohol consumption during pregnancy were prompted to state the advantages of abstinence. Moderately or not motivated women were asked to report on their perceived disadvantages of drinking during pregnancy. The midwife then advised them to stop drinking alcohol. In step 3, the barriers for successful abstinence and the mobilization of social support were discussed. In step 4, a self-help guide, adapted from an intervention on smoking in pregnancy [29], and relevant websites were mentioned. The midwife stimulated the pregnant woman to develop action plans for abstinence and coping with problems they might encounter when trying not to drink alcohol. If appropriate, access to alcohol addiction services was discussed. In step 5, women were asked to set a date for stopping their alcohol use (goal setting). Feedback session 2, approximately 8 weeks after baseline, consisted of step 6, which was addressed in approximately 1 minute (Feedback 2-health counseling). In this step, midwives again assessed the alcohol use of the pregnant women and asked her if she needed additional support for not drinking alcohol. Feedback session 3, approximately 14 weeks after baseline, consisted of step 7, which was also addressed in approximately 1 minute (Feedback 3-health counseling). In this step, midwives discussed alcohol use and its implications for breastfeeding.

Computer Tailoring

The computer-tailored intervention was developed using Tailorbuilder software (OSE, the Netherlands), a program which is specifically designed to develop Web-based computer-tailored interventions. Respondents in the computer-tailoring group received usual care from their midwife and computer-tailored feedback via the Internet, which was iterative and item-based [35]. Feedback 1, given immediately after baseline, consisted of 4 to 5 pages (Feedback 1-computer tailoring). This feedback was tailored to several respondent characteristics assessed in the baseline questionnaire: alcohol use, knowledge, risk perception, attitude, social influence, self-efficacy, intention, and action and coping plans. Specifically, the first feedback letter contained the recommendation of complete alcohol abstinence during pregnancy and information on possible consequences of prenatal alcohol use and the associated risk factors. In addition, feedback was provided on the respondent's risk perception of prenatal alcohol use; her attitude (perceived advantages and disadvantages toward prenatal alcohol use and alcohol abstinence; perceived social influence (not) to drink during pregnancy; self-efficacy to refrain from prenatal alcohol

XSL•FO
RenderX

use in specific situations, including suggestions on how to cope with these situations; the extent to which respondents were planning to undertake specific actions (action plans) to abstain from prenatal alcohol use; and how to cope with certain difficult situations (coping plans), including the formulation of personal plans in the shape of if-then statements [36]. The second feedback letter, 6 weeks after baseline, included personalized information on the respondents' choice of characteristics assessed with the baseline questionnaire (eg, risk perception or attitude; Feedback 2-computer tailoring). Depending on the number of characteristics chosen by the respondent, this feedback consisted of 1 or 2 pages. The third feedback letter, given immediately after T1, consisted of 3 to 4 pages of ipsative feedback tailored to changes in the respondent characteristics assessed at T1 in comparison to the baseline questionnaire (Feedback 3-computer tailoring). Feedback letters were visible on the computer screen and also sent to the respondent by email. Figure 1 shows an example of items regarding action plans to abstain from prenatal alcohol use. Figure 2 shows an example of a tailored feedback message.

Figure 1. Screenshot and English translation of items regarding action plans to abstain from prenatal alcohol use.



Figure 2. Screenshot and English translation of personal advice regarding prenatal alcohol use.



Usual Care

Midwives in the usual care group were instructed to give routine alcohol care. In-line with national guidelines, midwives recommend complete alcohol abstinence to clients who are using alcohol in the initial consultation [37,38]. In practice, not much time is spent on this nor is it common to provide additional counseling or other information [39].

Pretests of the Interventions

The midwives' manual of the health-counseling intervention was pretested among 5 midwives and the computer-tailoring intervention was pretested among 5 pregnant women using alcohol. The pretests yielded information about unclear questions and formulations in the manual and in the computer-tailoring intervention, which was used to improve the texts in the final versions of the health-counseling manual and computer-tailoring intervention.

Measures

Baseline Questionnaire

The baseline questionnaire required 15 minutes to complete, consisted of 92 questions, and was based on questionnaires in previous studies applying the I-Change model [40-42]. Questions assessed alcohol use in pregnancy (average alcohol use, binge drinking, and risky drinking), predisposing factors (drinking behavior before pregnancy, demographics, and smoking behavior), awareness factors (risk perception), and motivational factors (attitude, social influences, and self-efficacy).

Average alcohol consumption during pregnancy was assessed with the 5-item Dutch Quantity-Frequency-Variability (QFV) questionnaire [43]. Respondents selected the type of alcoholic drinks that they had consumed since the beginning of their pregnancy, such as beer, wine, or cocktails. Respondents were

XSL•FO
RenderX

asked to indicate how many working days (Monday to Thursday) on average they had consumed this type of alcohol since the beginning of their pregnancy. Additionally, they were asked to indicate the quantity (number of sips, glasses, or bottles) they had usually consumed of this type of alcohol on these occasions. Similar questions were asked concerning alcohol consumption during weekend days (Friday to Sunday). The average number of drinking working days multiplied by the average alcohol consumption per working day plus the average number of drinking weekend days multiplied by the average alcohol consumption per weekend day comprised the average weekly alcohol consumption during pregnancy.

We assessed 2 types of hazardous prenatal drinking behavior because previous research on alcohol use in pregnancy often used either of these types [44]. First, binge drinking in pregnancy was assessed by asking respondents if they ever had 4 or more standard glasses of alcohol (consisting of 10 grams of pure alcohol) on 1 day since they knew they were pregnant (0=no; 1=yes). Second, risky drinking during pregnancy was assessed with the validated T-ACE screening tool [45] (0=not risky drinking; 1=risky drinking).

Average alcohol consumption before pregnancy was also assessed with the QFV and was calculated similarly to the average weekly alcohol consumption during pregnancy.

Demographic information regarding age, education (primary school/basic vocational school, secondary vocational school/high school degree, higher vocational school/college degree/university degree), income (0.5 or less; 0.5-2; more than 2 times net Dutch median household income), and relationship status (0=no steady partner; 1=steady partner) was assessed. Pregnancy-related questions included number of weeks pregnant, number of prior pregnancies (0, 1, 2, or more than 2 prior pregnancies), and experience with complications in a previous pregnancy (0=no; 1=yes).

Respondents' smoking behavior (in number of cigarettes per day) during and before pregnancy was assessed ("On average, how many cigarettes do you smoke per day").

Risk perception was operationalized with perceived likelihood that the baby would experience harm ("If I drink alcohol in pregnancy, the chance that it damages my baby is..."; 1=very low; 5=very high) and perceived severity of that harm ("If I drink alcohol in pregnancy, the severity of the damage due to that alcohol is..."; 1=not serious; 5=very serious) resulting from alcohol use during pregnancy.

Attitude concerning alcohol use in pregnancy was assessed using 12 items (–2=disagree; 2=agree). A factor analysis using varimax rotation revealed 2 components: advantages (pros) and disadvantages (cons) of not drinking alcohol in pregnancy. Pros were assessed by 6 items (α=.75), such as "If I do not drink any alcohol in pregnancy, it is better for my baby's health." Cons of not drinking alcohol in pregnancy were assessed by another 6 items (α=.82), such as "If I do not drink any alcohol in pregnancy, I feel more tense."

Social support to abstain from alcohol in pregnancy was assessed with 3 items on a 5-point scale (α=.91) asking respondents whether they were supported by their partner, mother, and friends to abstain from alcohol in pregnancy, such as "My partner supports me not to drink alcohol in pregnancy" (–2=totally disagree; 2=totally agree).

Self-efficacy toward alcohol abstinence in pregnancy in social situations was assessed by 6 items on a 5-point scale (α=.90), such as "How easy it is for you to abstain from alcohol when your partner drinks alcohol" (–2=very difficult; 2=very easy).

Follow-Up Questionnaires

Posttest drinking behavior ("Have you had at least 1 sip of alcohol since the previous questionnaire?"; 0=no; 1=yes) and average weekly alcohol consumption since the previous questionnaire (assessed with the QFV) were assessed at T1 and at T2.

Analyses

The respondents who had had a miscarriage since the baseline were excluded from the analyses. The other respondents who did not complete the posttest questionnaire remained in the dataset and were considered as missing at random (MAR). Because respondents were nested in midwifery practices, all analyses were conducted using a mixed model analysis (SPSS v19).

To test whether conditions differed with regard to dropout, logistic mixed model analyses of dropout at T1 and T2 (0=no dropout; 1=dropout due to miscarriage, being unreachable, or being no longer interested to participate) were conducted with condition, age, education, steady partner, number of prior pregnancies, alcohol use before pregnancy, and smoking as independent variables.

To check for potentially confounding variables, univariate linear regressions with condition as predictor were performed and tested whether baseline characteristics of respondents differed between the 3 conditions.

In a set of multiple logistic mixed model analyses, we investigated the effect of condition in addition to the effect of covariates (concepts of the I-Change model) on posttest drinking behavior at T1 and T2 (0=not drinking; 1=still drinking). Significant interactions of covariates with condition were detected in a set of multiple logistic regression analyses conducted in a top-down procedure in which the least significant interaction, with $P>.05$, was omitted from a subsequent analysis. Significant main effects of covariates were also detected in a set of multiple logistic regression analyses conducted in a top-down procedure in which the least significant main effect, with $P>.05$, was omitted from a subsequent analysis. If there were no significant interaction effects with condition, we conducted a final multiple logistic regression with condition and the significant main effects of covariates and drinking behavior at T1 and T2 as outcome variable. If there were significant interaction effects with condition, we probed the interaction to understand the role of condition. Following Hayes and Matthes [46], we used the pick-a-point approach and tested whether condition was significant at 3 points on the moderator variable (1 standard deviation below average, average, and 1 standard deviation above average).

XSL•FO
RenderX

For respondents who were still drinking alcohol at T1 and T2, we tested the effect of condition in addition to the effect of covariates on the reduction of alcohol use. We performed similar sets of analyses as described previously using multiple linear mixed model analysis to assess the effect of condition in addition to the effect of confounding and moderating variables on average weekly alcohol consumption. Because of a right-skewed distribution (relatively few respondents had a high average weekly alcohol consumption), a transformation by the natural logarithm was applied to the average weekly alcohol consumption at T1 and T2.

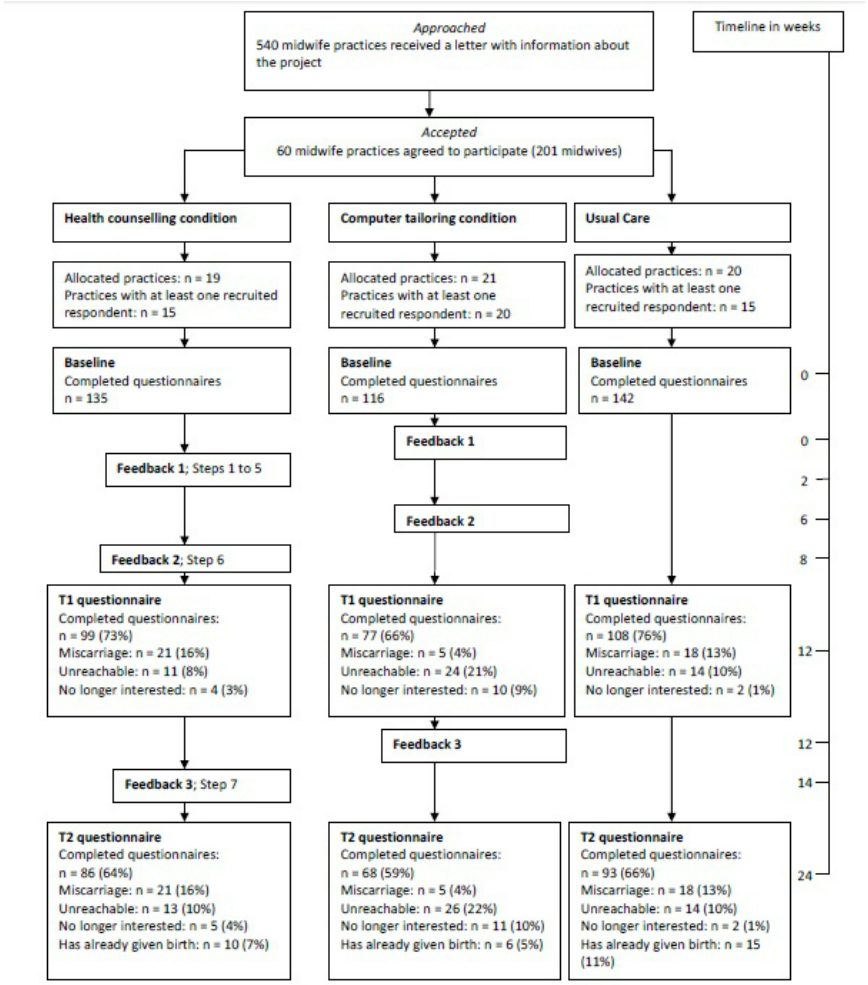
Finally, sensitivity analyses were conducted to test the robustness of the MAR assumption for the first hypothesis. These sensitivity analyses comprised the elaboration of 3 scenarios. First, all missing values were considered as still drinking alcohol; second, all missing values were considered as having stopped drinking alcohol. The third scenario entailed that women in the health-counseling condition who had quit alcohol were as likely as those who had not quit alcohol to return the follow-up questionnaire (eg, because of their connection with their midwife) whereas women in the computer tailoring and usual care conditions who had quit alcohol were twice as likely to return the follow-up questionnaire than those who had not quit alcohol (eg, because they wanted the researchers to know they had been successful). The robustness of the MAR assumption is supported when outcomes of these scenarios (including significant covariates) are similar to the outcomes of the analyses without the imputation of the missing values [47].

Results

Recruitment Results

The baseline questionnaire was completed by 393 respondents. In total, 135 respondents were assigned to the health-counseling condition, 116 respondents to the computer-tailoring condition, and 142 respondents to the usual care condition (Figure 3). These numbers varied slightly per condition because midwives in the 3 conditions yielded a slightly different number of participating women.

Figure 3. Flowchart of cluster randomized trial testing the effectiveness of health counseling and computer tailoring compared to usual care.



Selective Dropout

A total of 99 of 135 health-counseling respondents (73.3%), 77 of 116 computer-tailoring respondents (66.4%), and 108 of 142 usual care respondents (76.1%) completed T1. Multilevel logistic regression analysis with dropout (no/yes) at T1 as outcome variable and condition, age, education, steady partner, number of prior pregnancies, alcohol use before pregnancy, and smoking as covariates showed a significant fixed effect for having a steady partner. Respondents without a steady partner (OR 0.497, 95% CI 0.305-0.809; *P*=.005) were significantly

XSL•FO
RenderX

less likely to drop out at T1. Condition and random effects were not significant.

A total of 86 of 135 health-counseling respondents (63.7%), 68 of 116 computer-tailoring respondents (58.6%), and 93 of 142 usual care respondents (65.5%) completed the T2 questionnaire. Multilevel logistic regression analysis with dropout (no/yes) at T2 as outcome variable did not show a significant effect for condition or any other factor.

Sample Characteristics

Analyses on sample characteristics were conducted on the baseline characteristics of all respondents except 44 respondents with a miscarriage (114 health-counseling respondents, 111 computer-tailoring respondents, 124 usual care respondents; see Table 1). This sample had a mean age of 32.6 (SD 4.20) years. Most women were highly educated and had a medium income. The respondents were, on average, nearly 8 weeks pregnant, had been drinking almost 6 standard drinks of alcohol per week prior to pregnancy, and drank 1 standard drink of alcohol per week during pregnancy.

Multilevel analyses with sample characteristics as outcome variables and condition as predictor showed that computer-tailoring respondents drank significantly less alcohol before pregnancy compared to usual care respondents, and that health-counseling and computer-tailoring respondents smoked cigarettes more often compared to usual care respondents. Thus, alcohol use before pregnancy and smoking were considered potentially confounding variables in subsequent analyses. At the level of midwifery practices, respondents differed significantly with regard to educational level, income, and number of weeks of pregnancy indicating the necessity of a multilevel approach in the subsequent analyses.

Table 1. Baseline sample characteristics of Dutch pregnant women using alcohol (N=349).

Characteristic	Overall sample (N=349)	Health counseling (n=114)	Computer tailoring (n=111)	Usual care (n=124)	Condition effect, *P*	Random effect of midwifery practices, *P*
Age (years), mean (SD)	32.56 (4.20)	31.75 (4.37)	32.31 (4.22)	33.53 (3.85)	.17	.07
Educational level, n (%)					.15	.02
Low	9 (2.6)	5 (4.5)	1 (0.9)	3 (2.4)		
Medium	108 (31.2)	47 (42.0)	41 (36.9)	20 (16.3)		
High	229 (66.2)	60 (53.6)	69 (62.2)	100 (81.3)		
Income, n (%)					.93	.03
Low	35 (11.3)	14 (13.9)	9 (9.0)	12 (11.0)		
Medium	170 (54.8)	56 (55.4)	62 (62.0)	52 (47.7)		
High	105 (33.9)	31 (30.7)	29 (29.0)	45 (41.3)		
Steady partner, n (%)	198 (56.7)	73 (64.0)	66 (59.5)	59 (47.6)	.17	.33
Number of prior pregnancies, n (%)					.33	.14
0	150 (43.0)	51 (44.7)	37 (33.3)	62 (50.0)		
1	113 (32.4)	30 (26.3)	44 (39.6)	39 (31.5)		
2	56 (16.0)	23 (20.2)	16 (14.4)	17 (13.7)		
>2	30 (8.6)	10 (8.8)	14 (12.6)	6 (4.8)		
Number of weeks pregnant, mean (SD)	7.87 (1.96)	7.96 (1.81)	7.73 (2.06)	7.92 (1.99)	.72	.02
Experienced complications in previous pregnancy, n (%)	76 (22.4)	23 (20.9)	26 (23.9)	27 (22.3)	.87	.92
Standard alcohol drinks per week during pregnancy, mean (SD)	1.13 (2.87)	1.44 (3.33)	1.21 (3.14)	0.76 (2.02)	.23	.72
Binge drinkers during pregnancy,[a] n (%)	4 (1.2)	3 (2.7)	0 (0)	1 (0.8)	.17	—
Risky drinkers (T-ACE positive), n (%)	198 (57.4)	73 (64.6)	55 (50.9)	70 (56.5)	.13	.93
Standard alcohol drinks per week before pregnancy, mean (SD)	5.83 (7.35)	5.61 (8.88)	4.53 (4.61)[b]	7.18 (7.59)	.06	.62
Smokes in pregnancy,[a] n (%)	69 (20.2)	30 (27.0)[b]	25 (23.4)[b]	14 (11.3)	.01	—

[a] Single-level analyses were conducted on the characteristics binge drinking and smoking during pregnancy because in the multilevel analyses, the estimates of the variances of the random effects were 0 and the Hessian matrices were not positive definite.

[b]Indicates significant difference compared to usual care.

XSL•FO
RenderX

Drinking Behavior at T1 and T2

Our first hypothesis stated that women receiving health counseling or computer tailoring would be more likely to stop using alcohol in pregnancy compared to women receiving usual care at T1 and at T2. The results at T1 did not support our hypothesis. At T1, 64 of 99 health-counseling respondents (65%), 54 of 77 computer-tailoring respondents (70%), and 49 of 108 usual care respondents (45.4%) had refrained from alcohol. These differences were not significant (*P*=.79 for health counseling vs usual care; *P*=.15 for computer tailoring vs usual care; *P*=.23 for health counseling vs computer tailoring).

At T2, 62 of 86 health-counseling respondents (72%), 53 of 68 computer-tailoring respondents (78%), and 51 of 93 of the usual care respondents (55%) had refrained from alcohol. Table 2 presents the final model of the multilevel multiple logistic regression analyses with drinking behavior at T2 as outcome variable. The effects of the covariates (ie, alcohol use before pregnancy, smoking, age, education, perceived likelihood and perceived severity of risk due to prenatal alcohol use, pros and cons of not drinking alcohol in pregnancy, social support to abstain from alcohol in pregnancy, and social self-efficacy) were tested for significance and, if not significant, removed from the analysis model. The final analysis model showed that computer-tailoring respondents had refrained from alcohol significantly more often compared to usual care respondents, supporting our first hypothesis. However, the difference between health-counseling and usual care respondents was not significant (*P*=.26). Moreover, the difference between computer-tailoring and health-counseling respondents was not significant (*P*=.32).

Table 2. Final model of the multilevel multiple logistic regression analysis concerning drinking behavior at T2 (N=241).[a]

Fixed effects	Estimated variance	B	SE	OR	95% CI	*P*
Health counseling[b]		0.52	0.46	1.68	0.68, 4.18	.26
Computer tailoring[b]		1.02	0.49	2.77	1.05, 7.34	.04
Age		–0.11	0.05	0.89	0.82, 0.98	.01
Perceived likelihood		0.48	0.16	1.61	1.18, 2.19	.003
Self-efficacy		0.53	0.18	1.69	1.19, 2.41	.004
Random effect	0.40		0.36			.13

[a] 6 respondents were lost because they had not filled in the questions about self-efficacy in social situations.

[b] Usual care is the reference category.

Average Weekly Alcohol Consumption at T1 and T2

Our second hypothesis stated that women who continued their alcohol use would be more successful in reducing their alcohol consumption after receiving health counseling or computer tailoring at T1 and at T2. This hypothesis was not supported at T1. Only considering respondents who had not stopped drinking alcohol (n=35 in health counseling; n=23 in computer tailoring; n=59 in usual care), health-counseling respondents drank on average 0.56 standard drinks of alcohol per week (SD 0.91), computer-tailoring respondents drank 0.27 units (SD 0.17), and usual care respondents drank 0.51 units (SD 0.54). These differences were not significant (*P*=.58 for health counseling vs usual care; *P*=.23 for computer tailoring vs usual care; *P*=.49 for health counseling vs computer tailoring).

Our second hypothesis was partially supported at T2. Table 3 summarizes the results of the multilevel multiple linear regression analyses with average weekly alcohol consumption at T2 for those respondents who had not stopped drinking alcohol (n=23 in health counseling; n=15 in computer tailoring; n=41 in usual care). Health-counseling respondents drank on average 0.77 standard drinks of alcohol per week (SD 1.36), computer-tailoring respondents drank 0.35 units (SD 0.31), and usual care respondents drank 0.48 units (SD 0.54). Due to the significant interaction effects of computer tailoring with alcohol use before pregnancy and computer tailoring with social support, the main effect of computer tailoring could not be interpreted (see Table 3). To understand for which persons computer tailoring had a significant effect on the reduction of their alcohol use and for which persons computer tailoring was not significant, we probed the interactions by means of the pick-a-point approach. This showed that computer tailoring significantly reduced the alcohol use at T2 compared to usual care among respondents who had an average (*P*=.007) or 1 standard deviation below the average of alcohol use before pregnancy (*P*<.001), but not among respondents 1 standard deviation above the average of alcohol use before pregnancy (*P*=.57). In addition, computer tailoring significantly reduced alcohol use at T2 compared to usual care among respondents average (*P*=.03) or 1 standard deviation below the average of social support (*P*=.002), but not among respondents 1 standard deviation above the average of social support (*P*=.87). The analyses additionally showed that health counseling was not significant.

XSL•FO
RenderX

Table 3. Final model of the multilevel multiple linear regression analysis with the natural logarithm of average alcohol consumption at T2 as outcome variable among alcohol users only (N=73).[a]

Fixed effects	Estimated variance	B	SE	95% CI	*P*
Health counseling[b]		–1.11	0.92	–2.94, 0.72	.23
Computer tailoring[b]		6.41	1.75	2.92, 9.90	<.001
Not smoking[c]		–1.23	0.40	–2.03, –0.43	.003
Alcohol use before pregnancy		0.00	0.03	–0.06, 0.06	.95
Social support		0.16	0.16	–0.16, 0.47	.34
Health counseling[b] * alcohol use before pregnancy		–0.05	0.05	–0.15, 0.05	.32
Computer tailoring[b] * alcohol use before pregnancy		–0.43	0.12	–0.67, –0.18	.001
Health counseling[b] * social support		0.39	0.24	–0.08, 0.87	.10
Computer tailoring[b] * social support		–1.38	0.35	–2.08, –0.67	<.001
Random effect	0		0		

[a] 2 respondents were not included because they had not reported the amount of alcohol use; 6 respondents were lost because they had not filled in the question about social support.

[b] Usual care is the reference category.

[c] Smoking is the reference category.

Sensitivity Analyses for Missing-at-Random Assumption of Posttest Drinking Behavior (No/Yes)

We conducted sensitivity analyses for drinking behavior at T2 because with this outcome variable computer tailoring differed significantly from usual care. Scenario 1 entailed that all missing values were replaced with 1 (still drinking alcohol at T2). In scenario 1, 62 of 135 health-counseling respondents (45.9%), 53 of 116 computer-tailoring respondents (45.7%), and 51 of 142 usual care respondents (35.9%) refrained from alcohol. A multiple logistic regression analysis with drinking behavior 6 months after baseline according to scenario 1 as outcome variable showed that more computer-tailoring respondents refrained from alcohol than usual care respondents (*P*=.06). The difference between health counseling and usual care was not significant (*P*=.46).

Scenario 2 entailed that all missing values were replaced by 0 (not drinking alcohol at T2). In scenario 2, 111 of 135 health-counseling respondents (82.2%), 101 of 116 computer-tailoring respondents (87.1%), and 100 of 142 usual care respondents (70.4%) refrained from alcohol. A multiple logistic regression analysis with drinking behavior 6 months after baseline according to scenario 2 as outcome variable showed that significantly more computer-tailoring respondents refrained from alcohol than usual care respondents (*P*=.04). The difference between health counseling and usual care was not significant (*P*=.35).

In scenario 3, 36 of 49 missing values in the health-counseling condition (73%) were randomly replaced by 0 (not drinking alcohol at T2) and 13 of 49 missing values (27%) by 1 (still drinking alcohol at T2); in the computer-tailoring condition, 19 of 48 missing values (40%) were randomly replaced by 0 and 29 of 48 missing values (60%) by 1; in the usual care condition, 14 of 49 missing values (29%) were randomly replaced by 0 and 35 of 49 missing values (71%) by 1. A multiple logistic regression analysis with drinking behavior 6 months after baseline according to scenario 3 as outcome variable showed that significantly more computer-tailoring respondents (72/116, 62.1%) and health-counseling respondents (98/135, 72.6%) refrained from alcohol than usual care respondents (65/142, 45.8%; *P*=.04 and *P*=.01, respectively).

Discussion

The goal of this study was to test the effectiveness of 2 different brief interventions to reduce prenatal alcohol use, a health-counseling and a computer-tailored intervention, in comparison with usual care. We hypothesized that women receiving a newly developed health counseling or computer-tailored intervention were more likely to stop (hypothesis 1) and reduce (hypothesis 2) their prenatal alcohol use compared to women receiving usual care. This effect study showed that after 6 months and 3 feedback letters, the computer-tailoring program was effective in stopping prenatal alcohol use and in reducing it under certain conditions compared to usual care; the health-counseling protocol was not.

The ineffectiveness of the newly developed health-counseling protocol was inconsistent with the significant effects of health-counseling interventions in the related field of smoking cessation in pregnancy [29,48,49]. One shortcoming of this effectiveness study was the lack of statistical power. The power was planned to be .80 but turned out to be approximately .50 due to a larger intraclass correlation and a higher percentage of usual care participants who continued drinking than estimated beforehand. Although this amount of power was sufficient to show a significant effect of the computer-tailoring intervention at 6 months after baseline, the 20% difference between health-counseling and usual care respondents who stopped drinking alcohol at T1 and the 17% difference at T2 were not

XSL•FO
RenderX

significant. It is unclear whether health counseling would have been found effective with more statistical power.

A second shortcoming of this study was the suboptimal implementation of the health-counseling intervention by the midwives. Our process evaluation showed that the health-counseling midwives gave counseling less extensively than they were trained [50]. For example, the majority of midwives did not offer the second and third counseling sessions because they thought their clients did not need or like to receive this successive counseling. Also, in the related field of smoking cessation in pregnancy, results were ineffective when health professionals were found to lack skills to implement their tasks as intended [29,51]. A review on the barriers and facilitators of the effective implementation of brief interventions for alcohol misuse does show that effective implementation requires adequate training in which practitioners obtain sufficient confidence and knowledge to address drinking behavior without being worried to upset patients [52].

Both shortcomings imply that the ineffectiveness of the health-counseling intervention may not be simply due to an unsuccessful protocol. Perhaps the health-counseling protocol would have led to significant effects on the reduction of prenatal alcohol use with a higher amount of power and a better implementation. Future researchers testing an intervention to reduce prenatal alcohol use are recommended to take these issues into consideration (eg, [53]).

This is the first study showing that computer tailoring is effective in reducing prenatal alcohol use. The presently reported effect is in-line with previous studies showing how computer tailoring can effectively change health-related behaviors, such as smoking [40], vegetable and fruit intake [23], and alcohol use [24]. This computer-tailoring intervention is a promising method to reduce prenatal alcohol use. The high percentage of pregnant women using alcohol in the Netherlands [4] shows that alternatives to usual care are needed. Previous research has shown that pregnant women are reluctant to disclose their alcohol use to health professionals (eg, [39]). Because computer tailoring preserves a person's anonymity [54], computer tailoring may be an attractive intervention for these women. Moreover, the implementation of computer tailoring is not affected by barriers to the effective implementation of health counseling interventions, such as lack of resources, training, and support from management, as well as workload of practitioners providing health counseling [52]. Finally, previous research has shown that computer tailoring can be cheaper than a health-counseling intervention [55,56]; therefore, it may be a cost-effective method to decrease prenatal alcohol use, although additional research is needed to support this supposition.

A major strength of the present study was the use of a theoretical framework, which has been previously used in interventions for a variety of health behaviors (eg, [22,29,57]). In addition, both interventions used 3 feedback moments. Previous research on computer tailoring has shown that multiple feedback moments are likely to be more effective than a single feedback moment [58-60]. More research is needed to explore the optimal number of feedback moments for both computer tailoring and for health-counseling interventions. A limitation of the present study is the high percentage of dropout of respondents, especially in the computer-tailoring condition. Nevertheless, our sensitivity analyses show that the effectiveness of the computer-tailoring intervention is robust despite this high percentage of dropout. Another potential limitation is the reliance on self-report of alcohol use. Although the QFV is considered reasonably reliable [43], the use of more objective assessments, such as urine tests, may have yielded different results. Nevertheless, self-report methods of drinking (eg, QFV, the Alcohol Timeline Followback [61]) have been used in many studies on human drinking behavior because they are inexpensive, noninvasive, and acceptable to respondents [62]. Moreover, it is likely that the potential underreporting of alcohol use has occurred to an equal extent in the experimental and control conditions, upholding the effectiveness of computer tailoring. Finally, it was not possible to compare the effectiveness of computer tailoring with health counseling due to various differences in the set-ups of the interventions, including the anonymity of the respondents and the timing of the feedback. Only when the set-ups of the interventions are identical, future research will be able to compare the effectiveness of computer tailoring with health counseling.

To conclude, this research tested the effectiveness of 2 newly developed interventions to reduce prenatal alcohol use. Despite previous studies showing effects of health counseling in reducing prenatal alcohol use, our health-counseling intervention was not effective. Future studies testing health-counseling interventions are recommended to invest more in recruitment of pregnant women and implementation by health care providers. Our computer-tailoring intervention was effective in stopping and reducing prenatal alcohol use at 6-month follow-up. A cost-effectiveness study is recommended to determine the costs and effects associated with this intervention and compare them with the costs and effects of other interventions and/or usual care. A cost-effective computer-tailoring intervention would call for a broad implementation to prevent adverse neurodevelopmental effects in children due to light or moderate alcohol use.

Acknowledgments

This study was funded by a grant of the Dutch Organization for Health Research and Development (ZonMW). The authors would like to thank the midwifery practices and pregnant women for their participation. Preliminary results of this study were presented at the Second European Conference on FASD in Barcelona, Spain, in October 2012.

XSL•FO
RenderX

Conflicts of Interest

HdV is the scientific director of Vision2Health, a collaborating company between the University of Maastricht and OSE with the aim of offering proven effective methods in the field of health education. NvdW, CH, KE, MC, and WvD declare that they have no conflicts of interest.

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.2 [63].

[PDF File (Adobe PDF File), 1003KB - jmir_v16i12e274_app1.pdf]

References

1. From the Centers for Disease Control and Prevention. Alcohol use among women of childbearing age--United States, 1991-1999. JAMA 2002 Apr 24;287(16):2069-2071. [Medline: 11977236]
2. Nathanson V, Jayesinghe N, Roycroft G. Is it all right for women to drink small amounts of alcohol in pregnancy? No. BMJ 2007 Oct 27;335(7625):857 [FREE Full text] [doi: 10.1136/bmj.39356.489340.AD] [Medline: 17962287]
3. American Academy of Pediatrics Committee on Substance Abuse and Committee on Children with Disabilities: Fetal alcohol syndrome and fetal alcohol effects. Pediatrics 1993 May;91(5):1004-1006. [Medline: 8507280]
4. Health Council of the Netherlands. Risks of Alcohol Consumption Related to Conception, Pregnancy and Breastfeeding. Publication No. 2004/22. The Hague: Health Council of the Netherlands; 2005.
5. National Health and Medical Research Council. Australian Guidelines to Reduce Health Risks from Drinking Alcohol. Canberra: Commonwealth of Australia; 2009.
6. Chang G, Goetz MA, Wilkins-Haug L, Berman S. A brief intervention for prenatal alcohol use: an in-depth look. J Subst Abuse Treat 2000 Jun;18(4):365-369. [Medline: 10812310]
7. Gebara CF, Bhona FM, Ronzani TM, Lourenço LM, Noto AR. Brief intervention and decrease of alcohol consumption among women: a systematic review. Subst Abuse Treat Prev Policy 2013;8:31 [FREE Full text] [doi: 10.1186/1747-597X-8-31] [Medline: 24016074]
8. Nilsen P. Brief alcohol intervention to prevent drinking during pregnancy: an overview of research findings. Curr Opin Obstet Gynecol 2009 Dec;21(6):496-500. [doi: 10.1097/GCO.0b013e328332a74c] [Medline: 19797951]
9. Stade BC, Bailey C, Dzendoletas D, Sgro M, Dowswell T, Bennett D. Psychological and/or educational interventions for reducing alcohol consumption in pregnant women and women planning pregnancy. Cochrane Database Syst Rev 2009(2):CD004228 [FREE Full text] [doi: 10.1002/14651858.CD004228.pub2] [Medline: 19370597]
10. Chang G, Wilkins-Haug L, Berman S, Goetz MA. Brief intervention for alcohol use in pregnancy: a randomized trial. Addiction 1999 Oct;94(10):1499-1508. [Medline: 10790902]
11. Chang G, McNamara TK, Orav EJ, Wilkins-Haug L. Brief intervention for prenatal alcohol use: the role of drinking goal selection. J Subst Abuse Treat 2006 Dec;31(4):419-424. [doi: 10.1016/j.jsat.2006.05.016] [Medline: 17084796]
12. Handmaker NS, Miller WR, Manicke M. Findings of a pilot study of motivational interviewing with pregnant drinkers. J Stud Alcohol 1999 Mar;60(2):285-287. [Medline: 10091968]
13. O'Connor MJ, Whaley SE. Brief intervention for alcohol use by pregnant women. Am J Public Health 2007 Feb;97(2):252-258. [doi: 10.2105/AJPH.2005.077222] [Medline: 17194863]
14. Reynolds KD, Coombs DW, Lowe JB, Peterson PL, Gayoso E. Evaluation of a self-help program to reduce alcohol consumption among pregnant women. Int J Addict 1995 Mar;30(4):427-443. [Medline: 7607777]
15. Miller WR, Rollnick S. Motivational Interviewing: Preparing People to Change Addictive Behavior. New York: Guilford Press; 1991.
16. de Vries H, Brug J. Computer-tailored interventions motivating people to adopt health promoting behaviours: introduction to a new approach. Patient Educ Couns 1999 Feb;36(2):99-105. [Medline: 10223015]
17. Hawkins RP, Kreuter M, Resnicow K, Fishbein M, Dijkstra A. Understanding tailoring in communicating about health. Health Educ Res 2008 Jun;23(3):454-466 [FREE Full text] [doi: 10.1093/her/cyn004] [Medline: 18349033]
18. Kreuter M, Farrell D, Olevitch L, Brennan L. Tailoring Health Messages: Customizing Communication with Computer Technology. Mahwah, NJ: Lawrence Erlbaum Associates; 2000.
19. Dijkstra A. Working mechanisms of computer-tailored health education: evidence from smoking cessation. Health Educ Res 2005 Oct;20(5):527-539 [FREE Full text] [doi: 10.1093/her/cyh014] [Medline: 15701665]
20. Krebs P, Prochaska JO, Rossi JS. A meta-analysis of computer-tailored interventions for health behavior change. Prev Med 2010;51(3-4):214-221 [FREE Full text] [doi: 10.1016/j.ypmed.2010.06.004] [Medline: 20558196]
21. de Vries H, Kremers SP, Smeets T, Brug J, Eijmael K. The effectiveness of tailored feedback and action plans in an intervention addressing multiple health behaviors. Am J Health Promot 2008;22(6):417-425. [doi: 10.4278/ajhp.22.6.417] [Medline: 18677882]

XSL•FO
RenderX

22. Smit ES, de Vries H, Hoving C. Effectiveness of a Web-based multiple tailored smoking cessation program: a randomized controlled trial among Dutch adult smokers. J Med Internet Res 2012;14(3):e82 [FREE Full text] [doi: 10.2196/jmir.1812] [Medline: 22687887]
23. Oenema A, Tan F, Brug J. Short-term efficacy of a web-based computer-tailored nutrition intervention: main effects and mediators. Ann Behav Med 2005 Feb;29(1):54-63. [doi: 10.1207/s15324796abm2901_8] [Medline: 15677301]
24. Boon B, Risselada A, Huiberts A, Riper H, Smit F. Curbing alcohol use in male adults through computer generated personalized advice: randomized controlled trial. J Med Internet Res 2011;13(2):e43 [FREE Full text] [doi: 10.2196/jmir.1695] [Medline: 21719412]
25. Schulz DN, Candel MJ, Kremers SP, Reinwand DA, Jander A, de Vries H. Effects of a Web-based tailored intervention to reduce alcohol consumption in adults: randomized controlled trial. J Med Internet Res 2013;15(9):e206 [FREE Full text] [doi: 10.2196/jmir.2568] [Medline: 24045005]
26. Tzilos GK, Sokol RJ, Ondersma SJ. A randomized phase I trial of a brief computer-delivered intervention for alcohol use during pregnancy. J Womens Health (Larchmt) 2011 Oct;20(10):1517-1524 [FREE Full text] [doi: 10.1089/jwh.2011.2732] [Medline: 21823917]
27. Driscoll CD, Streissguth AP, Riley EP. Prenatal alcohol exposure: comparability of effects in humans and animal models. Neurotoxicol Teratol 1990;12(3):231-237. [Medline: 2196422]
28. Adams G, Gulliford MC, Ukoumunne OC, Eldridge S, Chinn S, Campbell MJ. Patterns of intra-cluster correlation from primary care research to inform study design and analysis. J Clin Epidemiol 2004 Aug;57(8):785-794. [doi: 10.1016/j.jclinepi.2003.12.013] [Medline: 15485730]
29. de Vries H, Bakker M, Mullen PD, van Breukelen G. The effects of smoking cessation counseling by midwives on Dutch pregnant women and their partners. Patient Educ Couns 2006 Oct;63(1-2):177-187. [doi: 10.1016/j.pec.2005.10.002] [Medline: 16406475]
30. de Vries H, Mudde A, Leijs I, Charlton A, Vartiainen E, Buijs G, et al. The European Smoking Prevention Framework Approach (EFSA): an example of integral prevention. Health Educ Res 2003 Oct;18(5):611-626 [FREE Full text] [Medline: 14572020]
31. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot 1997;12(1):38-48. [Medline: 10170434]
32. Ajzen I. From intentions to actions: A theory of planned behavior. In: Kuhl J, Beckmann J, editors. Action Control: From Cognition to Behavior. Berlin: Springer-Verlag; 1985:11-39.
33. Te Poel F, Bolman C, Reubsaet A, de Vries H. Efficacy of a single computer-tailored e-mail for smoking cessation: results after 6 months. Health Educ Res 2009 Dec;24(6):930-940 [FREE Full text] [doi: 10.1093/her/cyp036] [Medline: 19574405]
34. Elfeddali I, Bolman C, Candel MJ, Wiers RW, De Vries H. The role of self-efficacy, recovery self-efficacy, and preparatory planning in predicting short-term smoking relapse. Br J Health Psychol 2012 Feb;17(1):185-201. [doi: 10.1111/j.2044-8287.2011.02032.x] [Medline: 22107073]
35. Strecher VJ, Shiffman S, West R. Moderators and mediators of a web-based computer-tailored smoking cessation program among nicotine patch users. Nicotine Tob Res 2006 Dec;8 Suppl 1:S95-101. [Medline: 17491176]
36. van Osch L, Lechner L, Reubsaet A, Wigger S, de Vries H. Relapse prevention in a national smoking cessation contest: effects of coping planning. Br J Health Psychol 2008 Sep;13(Pt 3):525-535. [doi: 10.1348/135910707X224504] [Medline: 17631695]
37. De Jonge A. KNOV Standpoint Preconception Care. Bilthoven: Royal Dutch Organisation of Midwives; 2005.
38. De Boer J, Zeeman K. Antenatal Support. Recommendations for Support, Interaction and Information. Utrecht: Royal Organisation of Midwives; 2008.
39. van der Wulp NY, Hoving C, de Vries H. A qualitative investigation of alcohol use advice during pregnancy: experiences of Dutch midwives, pregnant women and their partners. Midwifery 2013 Nov;29(11):e89-e98. [doi: 10.1016/j.midw.2012.11.014] [Medline: 23434309]
40. Dijkstra A, De Vries H, Roijackers J. Long-term effectiveness of computer-generated tailored feedback in smoking cessation. Health Educ Res 1998 Jun;13(2):207-214 [FREE Full text] [Medline: 10181019]
41. Hoving C, Mudde AN, de Vries H. Effect of recruitment method and setting on the composition of samples consisting of adult smokers. Patient Educ Couns 2007 Jan;65(1):79-86. [doi: 10.1016/j.pec.2006.05.008] [Medline: 16872796]
42. Hoving C, Mudde AN, Dijk F, de Vries H. Effectiveness of a smoking cessation intervention in Dutch pharmacies and general practices. Health Education 2010 Jan 05;110(1):17-29. [doi: 10.1108/09654281011008726]
43. Lemmens P, Tan ES, Knibbe RA. Measuring quantity and frequency of drinking in a general population survey: a comparison of five indices. J Stud Alcohol 1992 Sep;53(5):476-486. [Medline: 1405641]
44. Jones TB, Bailey BA, Sokol RJ. Alcohol use in pregnancy: insights in screening and intervention for the clinician. Clin Obstet Gynecol 2013 Mar;56(1):114-123. [doi: 10.1097/GRF.0b013e31827957c0] [Medline: 23314712]
45. Sokol RJ, Martier SS, Ager JW. The T-ACE questions: practical prenatal detection of risk-drinking. Am J Obstet Gynecol 1989 Apr;160(4):863-8; discussion 868. [Medline: 2712118]
46. Hayes AF, Matthes J. Computational procedures for probing interactions in OLS and logistic regression: SPSS and SAS implementations. Behav Res Methods 2009 Aug;41(3):924-936. [doi: 10.3758/BRM.41.3.924] [Medline: 19587209]

XSL•FO
RenderX

47. Enders CK. Applied missing data analysis. In: Applied Missing Data Analysis (Methodology In The Social Sciences). New York: The Guilford Press; 2010.
48. Bolman C, de VH, van Breukelen G. Evaluation of a nurse-managed minimal-contact smoking cessation intervention for cardiac inpatients. Health Educ Res 2002 Feb;17(1):99-116 [FREE Full text] [Medline: 11888048]
49. Pieterse ME, Seydel ER, DeVries H, Mudde AN, Kok GJ. Effectiveness of a minimal contact smoking cessation program for Dutch general practitioners: a randomized controlled trial. Prev Med 2001 Feb;32(2):182-190. [doi: 10.1006/pmed.2000.0791] [Medline: 11162345]
50. van der Wulp NY, Hoving C, de Vries H. Dutch midwives' experiences with implementing health counselling to prevent prenatal alcohol use. J Clin Nurs 2014 Nov;23(21-22):3286-3289. [doi: 10.1111/jocn.12525] [Medline: 24393490]
51. DePue JD, Goldstein MG, Schilling A, Reiss P, Papandonatos G, Sciamanna C, et al. Dissemination of the AHCPR clinical practice guideline in community health centres. Tob Control 2002 Dec;11(4):329-335 [FREE Full text] [Medline: 12432158]
52. Johnson M, Jackson R, Guillaume L, Meier P, Goyder E. Barriers and facilitators to implementing screening and brief intervention for alcohol misuse: a systematic review of qualitative evidence. J Public Health (Oxf) 2011 Sep;33(3):412-421 [FREE Full text] [doi: 10.1093/pubmed/fdq095] [Medline: 21169370]
53. Wilson GB, McGovern R, Antony G, Cassidy P, Deverill M, Graybill E, et al. Brief intervention to reduce risky drinking in pregnancy: study protocol for a randomized controlled trial. Trials 2012;13:174 [FREE Full text] [doi: 10.1186/1745-6215-13-174] [Medline: 23006975]
54. Cline RJ, Haynes KM. Consumer health information seeking on the Internet: the state of the art. Health Educ Res 2001 Dec;16(6):671-692 [FREE Full text] [Medline: 11780707]
55. Smit F, Lokkerbol J, Riper H, Majo MC, Boon B, Blankers M. Modeling the cost-effectiveness of health care systems for alcohol use disorders: how implementation of eHealth interventions improves cost-effectiveness. J Med Internet Res 2011;13(3):e56 [FREE Full text] [doi: 10.2196/jmir.1694] [Medline: 21840836]
56. Smit ES, Evers SM, de Vries H, Hoving C. Cost-effectiveness and cost-utility of Internet-based computer tailoring for smoking cessation. J Med Internet Res 2013;15(3):e57 [FREE Full text] [doi: 10.2196/jmir.2059] [Medline: 23491820]
57. van Stralen MM, de Vries H, Mudde AN, Bolman C, Lechner L. The long-term efficacy of two computer-tailored physical activity interventions for older adults: main effects and mediators. Health Psychol 2011 Jul;30(4):442-452. [doi: 10.1037/a0023579] [Medline: 21639638]
58. Brug J, Glanz K, Van Assema P, Kok G, van Breukelen GJ. The impact of computer-tailored feedback and iterative feedback on fat, fruit, and vegetable intake. Health Educ Behav 1998 Aug;25(4):517-531. [Medline: 9690108]
59. Dijkstra A, De Vries H, Roijackers J, van Breukelen G. Tailoring information to enhance quitting in smokers with low motivation to quit: three basic efficacy questions. Health Psychol 1998 Nov;17(6):513-519. [Medline: 9848801]
60. Elfeddali I, Bolman C, Candel MJ, Wiers RW, de Vries H. Preventing smoking relapse via Web-based computer-tailored feedback: a randomized controlled trial. J Med Internet Res 2012;14(4):e109 [FREE Full text] [doi: 10.2196/jmir.2057] [Medline: 22903145]
61. Sobell LC, Sobell MB. Timeline followback: A technique for assessing self-reported ethanol consumption. In: Allen J, Litten RZ, editors. Measuring Alcohol Consumption: Psychosocial and Biological Methods. Totowa, NJ: Humana Press; 1992:41-72.
62. Del Boca FK, Darkes J. The validity of self-reports of alcohol consumption: state of the science and challenges for research. Addiction 2003 Dec;98 Suppl 2:1-12. [Medline: 14984237]
63. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

ICC: intraclass correlation coefficient
MAR: missing at random
QFV: Quantity-Frequency-Variability

XSL•FO
RenderX

Edited by G Eysenbach; submitted 27.04.14; peer-reviewed by N Johnson, S Haug; comments to author 17.07.14; revised version received 16.09.14; accepted 22.09.14; published 05.12.14.

Please cite as:
van der Wulp NY, Hoving C, Eijmael K, Candel MJJM, van Dalen W, De Vries H
Reducing Alcohol Use During Pregnancy Via Health Counseling by Midwives and Internet-Based Computer-Tailored Feedback: A Cluster Randomized Trial
J Med Internet Res 2014;16(12):e274
URL: http://www.jmir.org/2014/12/e274/
doi:10.2196/jmir.3493
PMID:25486675

©Nickie Y van der Wulp, Ciska Hoving, Kim Eijmael, Math JJM Candel, Wim van Dalen, Hein De Vries. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 05.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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

Automated Indexing of Internet Stories for Health Behavior Change: Weight Loss Attitude Pilot Study

Ramesh Manuvinakurike[1], MS; Wayne F Velicer[2], PhD; Timothy W Bickmore[1], PhD

[1]College of Computer and Information Science, Northeastern University, Boston, MA, United States

[2]Cancer Prevention Research Institute, University of Rhode Island, Kingston, RI, United States

Corresponding Author:
Timothy W Bickmore, PhD
College of Computer and Information Science
Northeastern University
WVH202
360 Hutington Ave
Boston, MA, 02115
United States
Phone: 1 617 373 5477
Fax: 1 617 812 2589
Email: bickmore@ccs.neu.edu

Abstract

Background: Automated health behavior change interventions show promise, but suffer from high attrition and disuse. The Internet abounds with thousands of personal narrative accounts of health behavior change that could not only provide useful information and motivation for others who are also trying to change, but an endless source of novel, entertaining stories that may keep participants more engaged than messages authored by interventionists.

Objective: Given a collection of relevant personal health behavior change stories gathered from the Internet, the aim of this study was to develop and evaluate an automated indexing algorithm that could select the best possible story to provide to a user to have the greatest possible impact on their attitudes toward changing a targeted health behavior, in this case weight loss.

Methods: An indexing algorithm was developed using features informed by theories from behavioral medicine together with text classification and machine learning techniques. The algorithm was trained using a crowdsourced dataset, then evaluated in a 2×2 between-subjects randomized pilot study. One factor compared the effects of participants reading 2 indexed stories vs 2 randomly selected stories, whereas the second factor compared the medium used to tell the stories: text or animated conversational agent. Outcome measures included changes in self-efficacy and decisional balance for weight loss before and after the stories were read.

Results: Participants were recruited from a crowdsourcing website (N=103; 53.4%, 55/103 female; mean age 35, SD 10.8 years; 65.0%, 67/103 precontemplation; 19.4%, 20/103 contemplation for weight loss). Participants who read indexed stories exhibited a significantly greater increase in self-efficacy for weight loss compared to the control group ($F_{1,107}$=5.5, P=.02). There were no significant effects of indexing on change in decisional balance ($F_{1,97}$=0.05, P=.83) and no significant effects of medium on change in self-efficacy ($F_{1,107}$=0.04, P=.84) or decisional balance ($F_{1,97}$=0.78, P=.38).

Conclusions: Personal stories of health behavior change can be harvested from the Internet and used directly and automatically in interventions to affect participant attitudes, such as self-efficacy for changing behavior. Such approaches have the potential to provide highly tailored interventions that maximize engagement and retention with minimal intervention development effort.

(J Med Internet Res 2014;16(12):e285) doi:10.2196/jmir.3702

KEYWORDS

behavioral medicine; natural language processing; animation; consumer health; health informatics; self-efficacy

XSL•FO
RenderX

Introduction

Background

The importance of health behavior change on public health has been well established, with lifestyle behaviors such as smoking, physical activity, and diet widely acknowledged to have major impacts on morbidity and mortality [1]. Automated health behavior change interventions have the promise of providing much greater reach and lower cost compared with interventions delivered by professionals. However, automated interventions—whether delivered through the Web, mobile devices, telephone via interactive voice response (IVR), or other media—can be costly and time-consuming to author or adapt [2]. In addition, tailoring intervention messages to user characteristics, such as age, race, gender, geographic locale, literacy level, self-efficacy, and stage of change, is known to be important and lead to more efficacious outcomes [3]. However, manually authoring messages for every combination of tailoring parameters makes development even more expensive and unwieldy despite the existence of automated tools to assist in the process [2].

In this paper, we explore a fundamentally new approach to developing automated health behavior change interventions by leveraging existing health message content on the Internet. Specifically, we aim to repurpose existing personal stories of health behavior change reported in blogs, bulletin board posts, and publicly available emails as the source messages for automated health behavior change systems. For blogs alone, more than a million posts are made each day on the Web, with the most common topic being stories of personal experience [4].

Personal stories of health behavior change represent a particularly powerful messaging mechanism. Heath information that is grounded in the personal experience of members of a community is more likely to be engaging and less likely to be dismissed [5]. Curated, edited, and packaged personal stories have also been used in several health behavior interventions. For example, Houston et al [6] videotaped stories told by individuals with hypertension and used these in a DVD-based intervention to help other patients with hypertension control their chronic condition. In a randomized clinical trial with 299 participants, those receiving the story-based intervention had significantly lower blood pressure after 3 months compared to a control group that received attention-control DVDs [6]. In addition, many individuals prefer getting certain types of health information via informal communication with others rather than from health professionals. A recent poll by the Pew Internet and American Life project reported that 43% of respondents preferred practical advice about day-to-day health situations from medical professionals, whereas a surprising 46% preferred to get this advice from other informal sources and sources such as family, friends, and fellow patients [7].

There are 3 primary challenges in our approach: (1) determining how to find and curate stories that provide helpful information or motivation (ie, filter out misinformation and stories that could potentially be counterproductive or even do harm), (2) determining how to select the best story to provide to a user at a given time (the “indexing” problem), and (3) determining the best medium and manner of telling a selected story to a user. Our central focus in this paper is on the second problem, but we also explore 2 different mediums in our summative evaluation to provide some preliminary information to address the third problem. Succinctly stated, the indexing problem is how to choose the best story from among potentially millions of possibilities to tell a given user at a given time to maximize their likelihood of achieving a specified health behavior criterion at some time in the future.

We focused our pilot work in this area on weight loss promotion. Overweight and obesity have reached epidemic proportions in the United States, with 68.5% of adults overweight or obese (body mass index [BMI] 25 or greater), 34.9% obese (BMI ≥30), and 6.5% extremely obese (BMI ≥40) [8]. These conditions have been implicated in a range of health conditions, including heart disease, stroke, type 2 diabetes mellitus, and some types of cancer [9], and the economic impact of obesity alone was estimated at US $99.2 billion in 1995 [10].

In the remainder of this paper, we describe the development and validation of a health behavior change story-indexing algorithm informed by theories of health behavior change and health communication. We then describe an evaluation study in which we assess the impact of automatically indexed personal stories of weight loss on psychological constructs from behavioral medicine; specifically, stage of change, self-efficacy, and decisional balance for weight loss.

Related Work in Story Indexing

Story indexing has its roots in early story understanding systems in which story indexes were hand-coded for intelligent retrieval and inference [11]. Early work on story understanding focused on information extraction to identify the stories that contained specific topics of interest [12]. More recently, researchers have employed statistical methods of text classification to collect and classify stories from large online datasets. For example, Gordon et al [4] created a classifier that was able to identify a million personal weblog entries from the Spinn3r dataset. Some researchers also explored automated story indexing for advice giving. Domeshek et al [13] developed a system that used 500 indexes to enable retrieval of stories to be used for social advice.

Related Work in Storytelling by Conversational Agents

Several prior studies have now demonstrated the efficacy of health interventions delivered by automated natural language dialog [14]. In particular, animated conversational agents that simulate face-to-face counseling sessions with a health professional have been shown to be well-accepted and a good vehicle for conveying the social dimensions of health messages [15,16], and have also been used to deliver stories to users in health care interventions [17,18].

Story Corpus

We collected 260 stories from websites including PatientsLikeMe [19], Experience Project [20], HealthTalkOnline [21], and About.com [22], where people frequently write about their experiences with health behavior change. Table 1 shows the places from which the stories were sourced. The stories were

XSL•FO
RenderX

hand-selected (by the first author) to be first-person accounts of successful attempts to work toward weight loss behavior change, evidenced by self-reported improvements in motivation, intention to change, or specific behaviors.

The stories did not go through any other filtering process other than the manual selection. The stories were relatively short, with a mean 327 (SD 211, range 44-1076) words. An excerpt from a story in the corpus:

> *My name is XXXX and I am 32 years old. Worst moment was seeing my son, pick up my bad eating habits. As a 1-year-old, he was having fries at McDonald's because that's what he saw me putting in my mouth. I read an article in my husband's health magazine about BMI and life expectancy; it said that being obese can shorten your life by 5 to 10 years. That was scary. I didn't want to miss out on any moments with my son just because of poor eating patterns. After that, every time I was tempted to have a fry, I'd think, "I could be here a little bit longer if I just eat better"...Weigh yourself twice a day every morning and night. It's the best way to stay on track. Exercise, even if it's only for 10 minutes. Anything is better than nothing. I also lift weights to boost my metabolism.*

Table 1. Sources of the stories and the number of stories selected from the sources.

Website (sources of stories)	Number of stories
Experience Project [20]	40
PatientsLikeMe [19]	40
HealthTalkOnline [21]	40
About.com [22]	23
Weight Loss Resources [23]	20
Yahoo! Lifestyle [24]	20
Blubberbusters [25]	20
3 Fat Chicks on a Diet! [26]	15
Weight Loss Success Stories [27]	12
Fitbie [28]	9
People [29]	8
Good Housekeeping [30]	5
Ladies' Home Journal [31]	5
Woman's Day [31]	3
Total	260

Story Indexing

Our indexing algorithm was designed to select the best story to tell a particular user at a particular time based on theories from behavioral medicine, health communication, and linguistics. From behavioral medicine, we included indexes based on information from the transtheoretical model (TTM; also known as the "stages of change" model) to select stories that help advance a user in their behavior change trajectory. From health communication, we included indexes based on tailoring theory to select stories that were as relevant as possible to the user's characteristics. From linguistics, we included indexes that selected stories based on assessments of story quality and positive affective tone. We describe each of these indexing features subsequently.

Indexing Features from the Transtheoretical Model of Health Behavior Change

The TTM of health behavior change posits that individuals change by progressing through a series of 5 well-defined stages, and that the messages and actions that are the most helpful to an individual at a given time are based on the stage they are in [32]. The 5 stages of change are precontemplation (people are not intending to take action in the foreseeable future), contemplation (people are intending to change soon), preparation (people are intending to take action in the immediate future), action (stage in which people have made specific modifications in their lifestyles), and maintenance (people are working to prevent relapse).

The TTM has been applied to a wide range of health behaviors, including weight loss [33]. In addition to the 5 stages, TTM defines 12 categories of specific behavior change activities ("processes of change") and posits that the processes that are most helpful to a given person at a given time depend on the individual's stage and the behavior being changed (Table 2). For example, consciousness raising (eg, researching the benefits of weight loss) is most helpful for individuals who are not yet considering change (precontemplation stage), whereas stimulus control (eg, putting a gym bag near the door as a reminder to exercise) is most helpful for those actively taking steps to change (action stage). Similarly, dramatic relief (eg, emotionally charged portrayals of the problems of overweight people) and environmental re-evaluation (eg, reflecting on the impact the

XSL•FO
RenderX

behavior has on the environment, such as society's food supply) are most useful for precontemplators (Figure 1 shows the relationships for weight loss).

The TTM also posits that self-efficacy (the degree of confidence an individual has to change their behavior, based on social cognitive theory [34]) and decisional balance (the perceived pros and cons of change) [35] covary in systematic ways with stage of change, but with the exact nature of the relationship depending on the particular behavior being changed. In general, self-efficacy and pros increase monotonically, whereas cons decrease monotonically as an individual progresses through the 5 stages of change.

The stage of change construct reflects information about an individual's intention to change their behavior in the future (precontemplation, contemplation, preparation) or the duration for which they have maintained change at a criterion level (action, maintenance). This specific temporal information rarely occurs in anecdotal stories of change and, even if it is explicitly or implicitly related in a story, it would require very sophisticated natural language processing to accurately extract. In comparison, specific behavior change activities (processes) are frequently mentioned in stories and are significantly easier to identify compared to information explicitly relating the author's stage of change. Thus, in our approach we use text classification methods to automatically identify processes of change in stories as features for indexing. When the system indexes the best story to tell for a given user, it uses information about the user's stage of change and the most relevant processes for that stage to index the best stories to tell.

Table 2. Processes of change used in indexing algorithm [32].

Processes of change	Description
1. Consciousness raising	Attempt to seek out information concerning their problem behavior
2. Dramatic relief	Increased emotional experiences followed by reduced affect if appropriate action can be taken
3. Substance use	Use of medication
4. Social liberation	Increase in social opportunities
5. Self–re-evaluation	Cognitive and affective assessments of one's self-image
6. Stimulus control	Removes cues for unhealthy habits and adds prompts for healthier alternatives
7. Helping relationship	Combine caring, trust, openness, and acceptance as well as support for the healthy behavior change
8. Counter conditioning	Learning of healthier behaviors that can substitute for problem behaviors
9. Reinforcement management	Consequences for taking steps in a particular direction
10. Self-liberation	Belief that one can change and the commitment and recommitment to act on that belief
11. Environmental re-evaluation	Affective and cognitive assessments of how the presence or absence of a personal habit affects one's social environment

Figure 1. Relation between stages and processes of change for weight loss [32].



Indexing Features From Tailoring Theory

Tailored health messages are those that are uniquely individualized to each person [36]. The most common approaches to tailoring involve selecting or adapting messages based on user characteristics, such as age, gender, race, and ethnicity. Message tailoring can be readily automated and computer-based tailored message interventions have been developed and evaluated. Computer-tailored messages have been shown to be significantly more effective at changing attitudes and behavior compared to nontailored messages in a number of studies [3]. Tailoring is thought to work by making messages more subjectively relevant and thus increasing the likelihood that individuals will process and internalize them [36].

In our approach, we used text classification methods to automatically identify demographic characteristics of authors based on the text of their stories, including age, gender, education level, and race. When the system indexed the best story to tell for a given user, it used information about the user's demographics to index the best stories to tell by preferring stories whose authors matched the user along these dimensions.

XSL•FO
RenderX

Affective Tone Indexing Feature

Stories that have an overall positive emotional tone are known to be more engaging and impactful than stories that are negative [37]; in addition, emotion can play a significant role in certain processes of change (eg, dramatic relief). Thus, we used text classification techniques to automatically compute the overall emotional tone ("sentiment") of a story, both as a process-specific feature and as a user-invariant quality metric to prefer stories that are positive.

Text Coherence Indexing Feature

Finally, we used an existing algorithm (Coh-Metrix) to automatically compute the coherence of a given story as a rough measure of writing quality [38]. The Coh-Metrix scores were calculated for stories using the online Coh-Metrix tool. This metric indicates how semantically meaningful, continuous, and understandable a story is and, like affective tone, is used as a user-invariant quality metric to prefer stories that are more coherent. Stories with higher Coh-Metrix scores were selected above the stories with lower scores only when the stories scored equally on all other indexes.

Indexing Algorithm

A given story may describe several behavior change activities that span more than 1 process of change; thus, it may be important to identify all processes of change mentioned in a story. Use of a multiclass classifier that combines information from 12 independent binary classifiers may seem to be a reasonable approach. However, analysis of the set of processes of change that are most useful for each stage of change for weight loss yields only 5 unique sets of processes that need to be identified, decreasing the complexity of the text classification problem from 12 classes to 5 (Table 3). In sum, we have developed a relevant stage of change classifier using the processes of change as features.

Table 3. Relevant stage of change classifier [19].

Class	Processes of change	Stage of change	Classification accuracy	F-measure
1	Consciousness raising, environmental re-evaluation, dramatic relief	Precontemplation	0.88	0.91
2	Consciousness raising, environmental re-evaluation, dramatic relief, helping relationship, social liberation	Contemplation	0.94	0.96
3	Helping relationship, social liberation	Preparation	0.87	0.93
4	Helping relationship, social liberation, self-liberation, self–re-evaluation, stimulus control, substance use, counter conditioning, reinforcement management	Action	0.98	0.95
5	Counter conditioning, reinforcement management	Maintenance	0.84	0.91

Indexing Algorithm Training and Accuracy Results

The stage of change classification algorithm training was performed using Adaptive Boosting (AdaBoost) [39] with weak learners using Weka [40]. Support vector machine, Naive Bayes, and regression classifiers were tried before settling with AdaBoost. The classifier producing the best accuracy without overfitting was selected for the task. The set of features used for classification included the presence of keywords for each process of change, sentiment scores of stories, part of speech (POS) tags, the position of keywords within a story, and any categorization tags associated with the story. Apache OpenNLP was used to obtain the parts of speech (eg, noun, verb) from the stories. The POS bigram and unigram counts in the story formed a set of features extracted. The keyword sets for each of the processes of change were formed from the stories. This feature was a binary valued indicator function. The value indicated if the story contained at least 1 keyword present in the keywords set defined for each process of change. As an example, the keywords for the substance use process of change (use of medication) included drugs, pills, medicine, tablets, etc. The indicator feature was generated to indicate the presence of substance use feature in the story. Similarly, the keywords for counterconditioning (learning of healthier behaviors that can substitute for problem behaviors) contained alter, change, substitute, lieu, etc. Helping relationship (combining caring, trust, openness, and acceptance as well as support for the healthy behavior change) keywords included mother, doctor, nurse, friend, help, motivate, etc. These indicator functions for each of the processes of change were constructed .and feature values were generated which indicated the presence of process of change keywords in the story. For each of the keyword lists, the position feature was constructed which indicated the part of the story (beginning, middle, or end). This function used 3 values which indicated if the keywords were present in the first, second, or the last third of the story based on the word count. For instance, if the keywords for a particular process of change was present in the first 100 words of a 300-word story, the feature took the value of "beginning." The same feature took the value of "end" if the substance use keyword was present in the last 100 words of the story. Stories containing dramatic relief processes of change generally contained higher sentiment scores. Numbers of noun and verb phrases were also used as a feature set obtained from Apache OpenNLP.

We trained our text classifiers on the corpus of 260 stories using calibration input from a crowdsourcing website. The calibration study was conducted using 260 participants on Amazon's Mechanical Turk. Northeastern University's institutional review board (IRB) approved the study. Participants each read a story from the corpus and answered 44 yes/no questions about it based on an existing processes of change questionnaire for weight loss [41]. These questions helped identify the processes of change from the stories by mapping each response to separate

XSL•FO
RenderX

processes of change. The responses to these questionnaires were mapped to the individual processes of change of TTM. The interrater reliability test was conducted on the mapped processes of change rather than the responses to the 44 questions. We conducted interrater reliability tests for the processes of change on 10 randomly selected stories by having 2 separate groups of users provide feedback. The resulting kappa was 0.80, indicating adequate reliability.

We used 10-fold cross-validation methodology to assess match accuracy on the story corpus. These values ranged from 83.8% to 98.1% for the 5 categories (Table 3).

We also constructed and trained text classifiers to automatically classify the gender of the story author using the same methodology as described previously. The age and race of the author were classified using regular expressions and their education level was computed using the story's readability score [38].

The final story-indexing algorithm used a user's demographic information (age, gender, education level, and race) and stage of change for weight loss (from a short questionnaire [35]) and scored each story in the corpus indicating the relative degree of relevance of the story to the user. The score was a weighted sum comprised of terms for degree of match between the user's demographic information and stage of change and the corresponding feature automatically computed from the story, as well as terms for the overall positive affect (sentiment) and quality (coherence) of the story. We weighted stage of change 5 times greater than the other terms in the score calculation.

Methods

Summary

We conducted an evaluation study to determine the impact of indexed stories on the health behavior change attitudes of individuals. This study was also conducted using a crowdsourcing site and was approved by Northeastern University's IRB. In this study, we sought to compare the impact of stories retrieved from the corpus by the indexing algorithm described previously and compared them to the impact of stories selected at random. We also sought to evaluate 2 different modalities for delivering the stories to individuals based on our prior work in using conversational agents for storytelling: an animated conversational agent vs text. Thus, our hypotheses were that for individuals interested in losing weight:

1. Indexed stories will have significantly greater impact than randomly selected stories on (1) changes in self-efficacy and decisional balance for weight loss and (2) ratings of story understandability, enjoyment, and identification with the story author.
2. Stories told by an animated conversational agent will have significantly greater impact than stories displayed in text on (1) changes in self-efficacy and decisional balance for weight loss and (2) ratings of story understandability, enjoyment, and identification with the story author.

To evaluate these hypotheses, we conducted a 2×2 full-factorial between-subjects experiment in which participants were randomized into 1 of 4 conditions according to 2 factors: (1) selection—stories were either selected using the indexing algorithm described previously (indexed) or selected at random (random) and (2) medium—stories were delivered either by a conversational agent (agent) or stories delivered in text (text).

Recruitment

The evaluation study was conducted using participants on Amazon's Mechanical Turk. Participants were required to be US-based "Master Turkers" (most experienced) and willing to lose weight.

Measures

The following self-report measures were assessed at enrollment and at the end of the intervention:

1. Stage of change for weight loss was assessed at enrollment and at the end of the intervention using the weight loss stages of change short form [33,35].
2. Weight loss self-efficacy was measured at enrollment and at the end of the intervention using a validated 10-item questionnaire [42].
3. Weight loss decisional balance was measured using a 20-item questionnaire [35].

In addition, the following measures were assessed at the end of the intervention:

1. Enjoyment of stories was measured using a single-item scale question ("How enjoyable was the story?").
2. Understandability of stories was assessed using items from a technology acceptance scale [43].

Identification with story was measured using 2 single-item scale questions ("How much can you identify with the story author?" and "How close were the experiences of the person in the story with yourself?").

Protocol

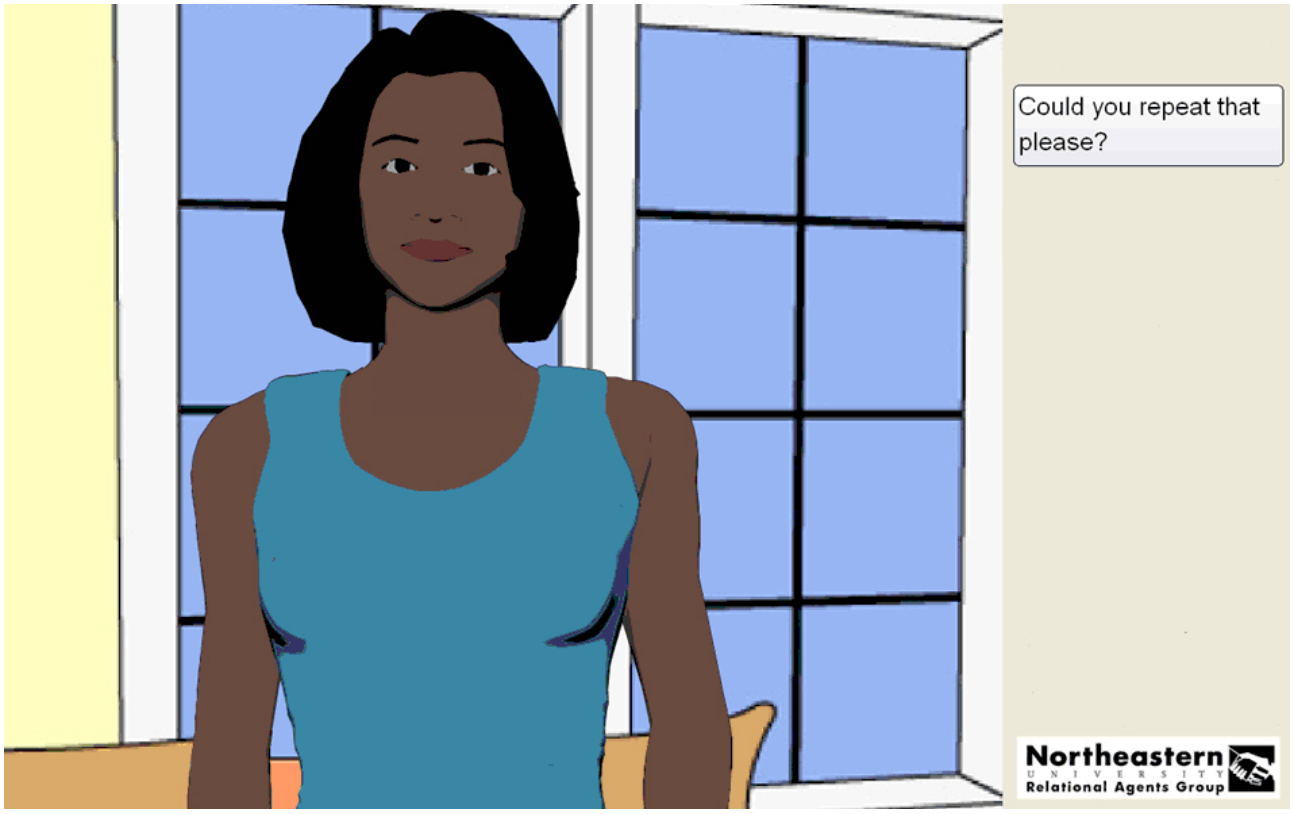
After passing our eligibility criteria and agreeing to an unsigned informed consent, participants filled out pretest questionnaires measuring demographics, stage of change, self-efficacy, and decisional balance.

Participants were then presented with the first story, either selected using the indexing algorithm (for indexed condition) or at random (for random condition). In text conditions, the story was simply displayed in a large text box. In agent conditions, an animated conversational agent was displayed that told the story verbally using synthetic speech and synchronized nonverbal conversational behavior (Figure 2) [44]. Participants were then presented with a questionnaire assessing enjoyment, understandability, and identification with the characters mentioned in the story.

Participants were then presented with a second story (in the same study treatment as the first) followed by a second administration of the questionnaire assessing enjoyment, understandability, and identification. Finally, participants were given a second administration of the self-efficacy and decisional balance questionnaires.

XSL•FO
RenderX

Figure 2. Animated conversational agent interface.



Statistical Analysis

Data were analyzed using 2×2 ANOVAs in SPSS (IBM Corp, Armonk, NY, USA) with selection and medium as the independent factors.

Results

Participants

Participant demographics are shown in Table 4. The 107 participants were recruited from 29 states in the United States. Only 4 of 107 participants withdrew from the study, all in the agent condition (2 in indexed, 2 in random). Therefore, 103 participants completed the entire study and were included in the final analysis; 53.4% (55/103) were female, the mean age was 34.85 (SD 10.8) years, and most were well educated (all had at least a high school education and 13.6%, 14/103 had graduate degrees).

Table 4. Participant demographics.

Measure	All participants	Agent		Text	
	(N=103)	Indexed (n=35)	Random (n=19)	Indexed (n=26)	Random (n=23)
Sex (female), n (%)	55 (53.4)	18 (51)	10 (53)	15 (58)	12 (52)
Age (years), mean (SD)	35.2 (11.1)	36.1 (12)	33.5 (11)	36.9 (10)	33.3 (11)
Race (white), n (%)	85 (82.5)	31 (89)	15 (79)	20 (77)	17 (74)
Education, n (%)					
<High school	0 (0.0)	0 (0)	0 (0)	0 (0)	0 (0)
High school	18 (17.5)	7 (20)	3 (16)	3 (12)	5 (22)
>High school	85 (82.5)	28 (80)	16 (84)	23 (89)	18 (78)
Stage of change, n (%)					
Precontemplation	67 (65.0)	24 (69)	12 (63)	17 (65)	14 (61)
Contemplation	20 (19.4)	5 (14)	5 (26)	5 (19)	5 (22)
Preparation	12 (11.7)	4 (11)	2 (11)	3 (12)	3 (13)
Action	4 (3.9)	2 (6)	0 (0)	1 (4)	1 (4)

XSL•FO
RenderX

Evaluation Outcomes

Table 5 summarizes results from the study. There were no significant effects of study manipulations on change in decisional balance for weight loss. There was a main effect of selection on change in self-efficacy for weight loss ($F_{1,107}$=5.5, P=.02) with indexed stories leading to significantly greater increases in self-efficacy compared to random stories. There was no effect of medium on change in self-efficacy ($F_{1,107}$=0.04, P=.84) and no significant interactions.

There was a significant main effect of medium on enjoyment of stories ($F_{1,98}$=22.3, P<.001), such that participants enjoyed reading stories themselves (text condition) significantly more compared to stories read by the agent (agent condition). There was no significant effect of selection on enjoyment.

There was a significant main effect of medium on identification with the story ($F_{1,98}$=47.2, P<.001), such that participants identified significantly more with stories in the text condition compared to those in the agent condition. There was also a trending main effect of selection on identification ($F_{1,98}$=3.0, P=.09) with participants identifying more with indexed stories compared to those selected at random.

Finally, there was a significant interaction of selection and medium on the rated understandability of stories ($F_{1,98}$=4.5, P=.04), such that participants receiving indexed stories rated those read by the agent as more understandable, whereas those who received random stories rated those delivered in text as more understandable. However, there was also a strong main effect of medium on understandability ($F_{1,98}$=82.8, P<.001) with participants in the agent group rating the understandability of their stories more highly overall compared to those in the text group.

Table 5. Primary outcomes by study condition.

Measure	Agent, mean (SD)		Text, mean (SD)	
	Indexed	Random	Indexed	Random
Self-efficacy (post-pre)	5.54 (4.54)	0.68 (7.40)	3.54 (3.66)	0.87 (3.98)
Decisional balance (post-pre)	0.37 (6.19)	1.89 (9.04)	0.27 (5.61)	–0.35 (2.23)
Enjoyment	2.15 (0.88)	2.45 (1.05)	3.15 (0.83)	2.89 (0.64)
Understandability	3.13 (0.85)	2.82 (0.99)	1.31 (0.69)	1.87 (0.64)
Identification	1.58 (0.73)	1.55 (0.71)	2.89 (0.86)	2.32 (0.63)

Discussion

Principal Results

Our first hypothesis (indexed stories will have significantly greater impact than randomly selected stories) received partial support, with indexed stories leading to significantly greater increases in self-efficacy and higher ratings of story understandability compared to stories selected at random. There was also a trend for indexed stories to lead to greater identification ratings. The lack of significant change in decisional balance could either be due to the complex relationship between self-efficacy and decisional balance (participants may not have been at the appropriate stage of change to see movement in both self-efficacy and decisional balance) or due to indexed stories acting uniquely on self-efficacy (confidence) without impacting decisional balance (perceived advantages or disadvantages of change).

However, our second hypothesis (stories told by an animated conversational agent will have significantly greater impact than stories displayed in text) received mixed support in the experiment. Although participants rated the understandability of stories more highly when read by an agent, they enjoyed them more and identified with them more when they read the stories from text. Medium had no effect on health behavior change attitudes.

Limitations

Relative to the millions of stories available on the Internet, the size of our experimental corpus of stories was tiny. However, we were able to demonstrate significant changes in health attitudes just with this small collection.

Our pool of participants was also very small and highly biased demographically. Although both genders were represented equally, Turkers are known to be younger, better educated, and earn more than the general US population [45].

Comparison With Prior Work

Houston et al [6] evaluated the impact of manually curated, edited, and packaged personal stories on health behavior change and demonstrated significant impacts on hypertension control compared to controls. However, this work is representative of the majority of research in behavioral medicine in which months or years are spent by experts developing and refining an intervention. In our approach, much of the intervention design is automated.

Lu [46] conducted a study in which she evaluated the effects of simulated blog posts on the health attitudes and intentions of individuals who read them. The posts were handcrafted to be either narrative (story) or nonnarrative, and to match study participants along various dimensions (“source similarity” which we have referred to as “tailoring”). Lu found that the tailoring effect was stronger in nonnarrative than narrative blogs. When the blogs were nonnarrative, those with health-related similarities were more effective at changing health-related

XSL•FO
RenderX

intentions than those with non-health-related similarities [46]. Although this study provides useful information that may improve our story-indexing approach, its objectives are fundamentally different from our goal of automatically selecting the best story to tell a given user from a collection of existing stories on the Internet.

Conclusions

We demonstrated the feasibility of automatically indexing stories of health behavior change gathered from the Internet and its ability to positively impact the health attitudes of individuals who read the stories.

We plan to continue improving our indexing algorithm with better features, improved machine learning methods, and a much larger training corpus. We also plan to investigate automating the process of identifying health behavior change stories on the Internet based on existing work on identifying personal stories in blog posts [4]. Automatically identifying and selecting stories that meet a particular individual's current needs and doing this in the context of a longitudinal health behavior change intervention represents a rich area of future research.

Acknowledgments

This work was supported by a grant from the National Institute on Minority Health and Health Disparities (R01MD006213). 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. Nandi A, Glymour MM, Subramanian SV. Association among socioeconomic status, health behaviors, and all-cause mortality in the United States. Epidemiology 2014 Mar;25(2):170-177. [doi: 10.1097/EDE.0000000000000038] [Medline: 24487200]
2. Vinson C, Bickmore T, Farrell D, Campbell M, An L, Saunders E, et al. Adapting research-tested computerized tailored interventions for broader dissemination and implementation. Transl Behav Med 2011 Mar;1(1):93-102 [FREE Full text] [doi: 10.1007/s13142-010-0008-9] [Medline: 24073035]
3. Noar SM, Benac CN, Harris MS. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. Psychol Bull 2007 Jul;133(4):673-693. [doi: 10.1037/0033-2909.133.4.673] [Medline: 17592961]
4. Gordon A, Swanson R. Identifying Personal Stories in Millions of Weblog Entries. 2009 Presented at: Third International Conference on Weblogs, Social Media, Data Challenge Workshop; May 17, 2009; San Jose, CA URL: http://www.icwsm.org/2009/data/Gordon-icwsm09-dcw.pdf
5. Hinyard LJ, Kreuter MW. Using narrative communication as a tool for health behavior change: a conceptual, theoretical, and empirical overview. Health Educ Behav 2007 Oct;34(5):777-792. [doi: 10.1177/1090198106291963] [Medline: 17200094]
6. Houston TK, Allison JJ, Sussman M, Horn W, Holt CL, Trobaugh J, et al. Culturally appropriate storytelling to improve blood pressure: a randomized trial. Ann Intern Med 2011 Jan 18;154(2):77-84. [doi: 10.7326/0003-4819-154-2-201101180-00004] [Medline: 21242364]
7. Fox S. The Social Life of Health Information. Washington, DC: Pew Internet & American Life Project; 2011 May 12. URL: http://www.pewinternet.org/2011/05/12/the-social-life-of-health-information-2011/# [accessed 2014-11-26] [WebCite Cache ID 6UNjrVOfv]
8. Flegal K, Carroll M, Kit B, Ogden C. Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999-2010. JAMA 2012 Feb 1;307(5):491-497. [doi: 10.1001/jama.2012.39] [Medline: 22253363]
9. Pi-Sunyer FX. Medical hazards of obesity. Ann Intern Med 1993 Oct 1;119(7 Pt 2):655-660. [Medline: 8363192]
10. Wolf AM, Colditz GA. Current estimates of the economic cost of obesity in the United States. Obes Res 1998 Mar;6(2):97-106. [Medline: 9545015]
11. Meuller E. Story understanding. In: Nadel L, editor. Encyclopedia of Cognitive Science. New York: Wiley; 2005:238-246.
12. Riloff E. Information extraction as a stepping stone toward story understanding. In: Ram A, Moorman K, editors. Understanding Language Understanding: Computational Models of Reading. Cambridge, MA: MIT Press; 1999.
13. Domeshek E. Do the Right Thing: A Component Theory for Indexing Stories as Social Advice [dissertation]. New Haven, CT: Yale University; 1992.
14. Bickmore T, Giorgino T. Health dialog systems for patients and consumers. J Biomed Inform 2006 Oct;39(5):556-571 [FREE Full text] [doi: 10.1016/j.jbi.2005.12.004] [Medline: 16464643]
15. Bickmore T, Gruber A, Picard R. Establishing the computer-patient working alliance in automated health behavior change interventions. Patient Educ Couns 2005 Oct;59(1):21-30. [doi: 10.1016/j.pec.2004.09.008] [Medline: 16198215]

XSL•FO
RenderX

16. Bickmore TW, Pfeifer LM, Byron D, Forsythe S, Henault LE, Jack BW, et al. Usability of conversational agents by patients with inadequate health literacy: evidence from two clinical trials. J Health Commun 2010;15 Suppl 2:197-210. [doi: 10.1080/10810730.2010.499991] [Medline: 20845204]
17. Bickmore T, Ring L. Making it personal: end-user authoring of health narratives delivered by virtual agents. In: Intelligent Virtual Agents Conference. 2010 Presented at: Intelligent Virtual Agents conference (IVA); September 20, 2010; Philadelphia, PA URL: http://relationalagents.com/publications/IVA10.story.pdf
18. Gardiner P, Hempstead MB, Ring L, Bickmore T, Yinusa-Nyahkoon L, Tran H, et al. Reaching women through health information technology: the Gabby preconception care system. Am J Health Promot 2013;27(3 Suppl):eS11-eS20 [FREE Full text] [doi: 10.4278/ajhp.1200113-QUAN-18] [Medline: 23286652]
19. HealthTalkOnline. URL: http://www.healthtalk.org/ [accessed 2014-11-21] [WebCite Cache ID 6UGIlRSTi]
20. PatientsLikeMe. URL: http://www.patientslikeme.com/ [accessed 2014-11-21] [WebCite Cache ID 6UGIkEy3G]
21. About.com. URL: http://www.about.com/ [accessed 2014-11-21] [WebCite Cache ID 6UGIgn127]
22. Weight Loss Resources. URL: http://www.weightlossresources.co.uk/ [accessed 2014-11-21] [WebCite Cache ID 6UGIZwvbu]
23. Yahoo! Lifestyle. URL: https://au.lifestyle.yahoo.com/ [accessed 2014-11-21] [WebCite Cache ID 6UGIWDcX5]
24. Blubberbusters. URL: http://www.blubberbuster.com/ [accessed 2014-11-21] [WebCite Cache ID 6UGIT2LmD]
25. 3 Fat Chicks on a Diet!. URL: http://www.3fatchicks.com/forum/ [accessed 2014-11-21] [WebCite Cache ID 6UGIRN9zx]
26. Weight Loss Success Stories. URL: http://weight-loss-story.blogspot.com/ [accessed 2014-11-21] [WebCite Cache ID 6UGIK6o2A]
27. Fitbie. URL: http://www.fitbie.com/ [accessed 2014-11-21] [WebCite Cache ID 6UGICWy13]
28. People. URL: http://www.people.com/people/ [accessed 2014-11-21] [WebCite Cache ID 6UGI9IqUl]
29. Good Housekeeping. URL: http://www.goodhousekeeping.com/ [accessed 2014-11-21] [WebCite Cache ID 6UGI7RUIK]
30. Ladies' Home Journal. URL: http://www.divinecaroline.com/ladies-home-journal [accessed 2014-11-21] [WebCite Cache ID 6UGI5zgRL]
31. Woman's Day. URL: http://www.womansday.com/ [accessed 2014-11-21] [WebCite Cache ID 6UGI4HsKH]
32. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot 1997;12(1):38-48. [Medline: 10170434]
33. Prochaska JO, Velicer WF, Rossi JS, Goldstein MG, Marcus BH, Rakowski W, et al. Stages of change and decisional balance for 12 problem behaviors. Health Psychol 1994 Jan;13(1):39-46. [Medline: 8168470]
34. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977 Mar;84(2):191-215. [Medline: 847061]
35. O'Connell D, Velicer W. A decisional balance measure and the stages of change model for weight loss. Int J Addict 1988 Jul;23(7):729-750. [Medline: 3192340]
36. Hawkins RP, Kreuter M, Resnicow K, Fishbein M, Dijkstra A. Understanding tailoring in communicating about health. Health Educ Res 2008 Jun;23(3):454-466 [FREE Full text] [doi: 10.1093/her/cyn004] [Medline: 18349033]
37. Pennebaker JW, Seagal JD. Forming a story: the health benefits of narrative. J Clin Psychol 1999 Oct;55(10):1243-1254. [doi: 10.1002/(SICI)1097-4679(199910)55:10<1243::AID-JCLP6>3.0.CO;2-N] [Medline: 11045774]
38. McNamara D, Graesser A, McCarthy P, Cai Z. Automated Evaluation of Text and Discourse With Coh-Metrix. Cambridge, UK: Cambridge University Press; 2014.
39. Freund Y, Schapire R. A Short Introduction to Boosting. Journal of Japanese Society for Artificial Intelligence 1999;14(5):771-780 [FREE Full text]
40. Holmes G, Donkin A, Witten I. Weka: a machine learning workbench. 1994 Presented at: Second Australian and New Zealand Conference on Intelligent Information Systems; Noveber 29- December 2, 1994; Brisbane, Australia URL: http://www.cs.waikato.ac.nz/~ihw/papers/94GH-AD-IHW-WEKA.pdf
41. Rossi S, Rossi J. Processes of change for dietary fat reduction. 1993 Presented at: Annual Meeting of the Society for Behavioral Medicine; March, 1993; San Francisco, CA.
42. 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]
43. Adams B, Nelson R, Hall M, Todd P. Perceived usefulness, ease of use, and usage of information technology: a replication. MIS Quarterly 1992;16(2):227-248.
44. Bickmore T, Schulman D, Shaw G. DTask and LiteBody: open source, standards-based tools for building web-deployed embodied conversational agents. : Springer; 2009 Presented at: Intelligent Virtual Agents (IVA); September 14, 2009; Amsterdam, Netherlands URL: http://relationalagents.com/publications/IVA09.litebody.pdf
45. Ross J, Irani L, Silberman M, Zaldivar A, Tomlinson B. Who are the crowdworkers?: shifting demographics in mechanical turk. In: CHI '10 Extended Abstracts on Human Factors in Computing Systems 2010. 2010 Presented at: Human Factors in Computing Systems (CHI); April 10, 2010; Atlanta, GA p. 2863-2872.
46. Lu AS. An experimental test of the persuasive effect of source similarity in narrative and nonnarrative health blogs. J Med Internet Res 2013;15(7):e142 [FREE Full text] [doi: 10.2196/jmir.2386] [Medline: 23887135]

XSL•FO
RenderX

Abbreviations

BMI: body mass index
IVR: interactive voice response
POS: part of speech
TTM: transtheoretical model

Edited by G Eysenbach; submitted 12.07.14; peer-reviewed by B Yu, L Toldo; comments to author 07.08.14; revised version received 12.10.14; accepted 06.11.14; published 09.12.14.

Please cite as:
Manuvinakurike R, Velicer WF, Bickmore TW
Automated Indexing of Internet Stories for Health Behavior Change: Weight Loss Attitude Pilot Study
J Med Internet Res 2014;16(12):e285
URL: http://www.jmir.org/2014/12/e285/
doi:10.2196/jmir.3702
PMID:25491243

©Ramesh Manuvinakurike, Wayne F Velicer, Timothy W Bickmore. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 09.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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 Mobile Apps for Health Promotion in Australian Indigenous Populations: Scoping Review

Carl Brusse[1], BSc, BA (Hons), MA; Karen Gardner[1], BA, MPH, PhD; Daniel McAullay[1], BSc, MAE, PhD; Michelle Dowden[2], RN, MPH

[1]Australian Primary Health Care Research Institute, Research School of Population Health, Australian National University, Canberra, Australia
[2]Independant consultant, Darwin, Australia

Corresponding Author:
Karen Gardner, BA, MPH, PhD
Australian Primary Health Care Research Institute
Research School of Population Health
Australian National University
Building 63
Mills & Eggleston Roads
Canberra, 2601
Australia
Phone: 61 6125 7875
Fax: 61 6230 0525
Email: karen.gardner@anu.edu.au

Abstract

Background: Health promotion organizations are increasingly embracing social media technologies to engage end users in a more interactive way and to widely disseminate their messages with the aim of improving health outcomes. However, such technologies are still in their early stages of development and, thus, evidence of their efficacy is limited.

Objective: The study aimed to provide a current overview of the evidence surrounding consumer-use social media and mobile software apps for health promotion interventions, with a particular focus on the Australian context and on health promotion targeted toward an Indigenous audience. Specifically, our research questions were: (1) What is the peer-reviewed evidence of benefit for social media and mobile technologies used in health promotion, intervention, self-management, and health service delivery, with regard to smoking cessation, sexual health, and otitis media? and (2) What social media and mobile software have been used in Indigenous-focused health promotion interventions in Australia with respect to smoking cessation, sexual health, or otitis media, and what is the evidence of their effectiveness and benefit?

Methods: We conducted a scoping study of peer-reviewed evidence for the effectiveness of social media and mobile technologies in health promotion (globally) with respect to smoking cessation, sexual health, and otitis media. A scoping review was also conducted for Australian uses of social media to reach Indigenous Australians and mobile apps produced by Australian health bodies, again with respect to these three areas.

Results: The review identified 17 intervention studies and seven systematic reviews that met inclusion criteria, which showed limited evidence of benefit from these interventions. We also found five Australian projects with significant social media health components targeting the Indigenous Australian population for health promotion purposes, and four mobile software apps that met inclusion criteria. No evidence of benefit was found for these projects.

Conclusions: Although social media technologies have the unique capacity to reach Indigenous Australians as well as other underserved populations because of their wide and instant disseminability, evidence of their capacity to do so is limited. Current interventions are neither evidence-based nor widely adopted. Health promotion organizations need to gain a more thorough understanding of their technologies, who engages with them, why they engage with them, and how, in order to be able to create successful social media projects.

(J Med Internet Res 2014;16(12):e280) doi:10.2196/jmir.3614

KEYWORDS

health promotion; indigenous health; eHealth

XSL•FO
RenderX

Introduction

Health Promotion, Health Outcomes, and Technology

One key strategy used in health promotion interventions is to engage a target audience among the general public, with the goal of improving health outcomes by influencing the behavior of that audience. The intended health concerns and behavioral goals will vary from project to project and the target audience might be a narrow demographic or broader and more general, but the basic nature of this particular aspect of health promotion is the same: a non-commercial marketing exercise in the name of a greater social good. This subset of health promotion can therefore be seen as a species of "social marketing".

Understood as social marketing, health promotion interventions can be, should be, and indeed are being taken beyond traditional public service messages and familiar advertising campaigns [1-3]. Marketing of all forms has become increasingly sophisticated in the online era, with new media modalities and analytics enabling more targeted approaches to both commercial and social marketing. Among the numerous health promotion projects and engagement strategies being explored, new media technologies such as social media and mobile phone/tablet apps have attracted a great deal of interest from public health organizations and researchers [3,4].

The potential for using new media to achieve improved health outcomes is broader than social marketing alone. The interactive nature and the personal intimacy of these new technologies for an "always online" health consumer allows for their use in self-management and even in the delivery of some health services. However, in part because these techniques are in their infancy (and for other reasons that shall be discussed), early evidence of effectiveness and benefit is lacking.

The research reported here is intended to provide a current overview of the evidence surrounding *consumer-use* social media and mobile software (apps) for health promotion interventions, with a particular focus on the Australian context and on health promotion targeted toward an Indigenous audience. This intent and focus directed three component scoping surveys: (1) a survey of the uses of social media to reach Indigenous Australians for health promotion, (2) a survey of Australian health promotion/social marketing projects that make use of social media and mobile apps, and (3) a scoping study of peer-reviewed evidence for the effectiveness of social media and mobile technologies in health promotion, globally.

The Internet and Social Media

It has been 10 years since the average time that young Americans spent online exceeded the time they spent watching TV; a statistic that the general population in developed countries has gradually been catching up with. The time spent on online social media has ballooned since the introduction of dedicated social media platforms in 2004-2007 and now amounts to at least 20% of all time spent online, with that proportion again much higher in the youth demographic. Publically available Web traffic measurements now rank the leading social media domain, Facebook.com, as either the first or second most popular domain in Australia, marginally ahead of or behind Google.com domains, depending on the methodology used. Indeed, some industry watchers have even speculated that the rapidly increasing proportion of online content being created, shared, and consumed within social media "walled gardens" may be altering the fundamental character of the Web, its openness, and how it is used.

For the moment though, and for the purposes of social marketing, including health promotion, the upshot of these changes is simple: social media is increasingly where the eyeballs are (not to mention fingertips). If there is a message to present or content to deliver, then establishing a social media presence represents both an opportunity and increasingly a requirement.

Social media also holds attractions for the social marketer beyond mere reach. For example, minimal resources are required to create a basic online presence in social media platforms such as Facebook: a Facebook page-making platform can generate a ready receptacle for publishing or publicizing updatable content, from simple comments and posts to images, video, and interactive media. The most salient and alluring promises of social media however (for health promotion and other social marketing) come with the social aspects of social media.

The "social" in social media refers to method and modality, rather than to goals or objectives as in "social marketing". The activities that constitute social media—users forming social connections and carrying out social interactions via the medium of social media platforms—can be seen as the extension of natural human social networking into online spaces. Social media users add or re-transmit pictorial, textual, or multimedia content that they find interesting, sharing it with peers, circles of contacts, or the Web at large. Social media therefore combines the reach of mass media, but is integrated into and driven by the small-scale social networking of everyday people. The Holy Grail for any social marketer is for their message to be picked up and transmitted widely via online social networks (to "go viral") and thereby radically magnify its reach and effect.

But social media is complex. First, because it is driven by user interest, it is up to the users to decide whether any particular content is taken up and re-transmitted (and with what framing, eg, serious or sarcastic). Second, the modalities of social media are themselves complex. Some social media is carried out via secondary features on specialized content-focused platforms (eg, YouTube.com), some within rich and freestanding feature-packages (such as with Facebook.com), and some via minimalist messaging and micro-blogging services (eg, Twitter.com). And the social media "workflow" of users can involve the mixing and matching of social media functions, online content, and other messaging and publishing tools (such as instant messaging/chat, Web forums, blogs, webpage comments, and even email) within a sophisticated ecosystem of online social behavior. Broad characterizations and analyses of social media usage must therefore be treated with caution, because social media is defined by use and function rather than by form or design.

For the health promotion practitioner, social media is therefore a complex beast with great potential for both success and failure—a way of starting and carrying on an actual conversation

XSL•FO
RenderX

with the target audience (as long as their attention is acquired and retained), which has the potential to take on a life of its own (for better or for worse) or simply be ignored.

Mobile Technology

A second major trend in recent online behavior is that more and more "eyeball time" takes place on mobile devices outside the Web browser. In 2011 (only 4 years after the introduction of the first iPhone), sales of mobile phones overtook those of traditional desktop and laptop computers and the expectation is that tablet devices will do the same within 2 years [5]. The rapid uptake of these mobile devices has driven a whole new class of lightweight software—mobile apps—to suit compact, touch-screen interfaces and to integrate with telephony, GPS, and sensor features included in the always-on, battery-dependent mobile hardware.

Well-written mobile apps are, in effect, pre-packaged caches of interactive content, installed on devices that users carry about with them. Whereas traditional websites for example require a constant, reliable Internet connection to be accessed, users have mobile apps on their phone at all times, including when Internet access is spotty or when they've run out of data.

Mobile apps are therefore a tantalizing target for health promotion purposes—and for similar reasons to social media. The uptake of social media and mobile apps is user-driven: meaning they are interventions that can potentially be disseminated much more widely than a traditional intervention where subjects are individually recruited. Furthermore, the "buy-in" effect is potentially very significant for apps that successfully make it onto mobile devices and actually get used. As with social media projects then, the trick would be to design apps that are attractive and accessible to the target population, as well as constituting effective interventions.

Indigenous Use of Social Media and Mobile Technology in Australia

For the purposes of our review, the target population of interest is Indigenous Australians: a traditionally underserved population with a high prevalence of public health issues, with particular areas of concern such as smoking cessation, sexual health, and otitis media commonly identified by health care workers.

International studies have suggested that mobile devices can help traditionally underserved populations by leapfrogging economic and infrastructural bottlenecks—bringing connectivity to individuals in remote or underserved communities without having to wait for access to more traditional (and expensive) residential Internet or computer use [3]. But Indigenous cultures in Australia also deeply embed practices of social relatedness, kinship, and identity with specific priorities for communication and social interconnection despite the remoteness of some of these communities. As might be expected then, there is a high penetration of mobile phone usage among Indigenous Australians including those in rural and remote communities. Mobile phones are by far the most used technology among adolescents [6], surpassing TV, video games, and other forms of Internet access. In the past 5 years, affordable mobile phones with camera and messaging functions have spawned a "mobile phone culture" in some remote areas, where messages, pictures, and video clips flow freely among and between communities, often in culturally unique and creative ways [7].

The questions for health promotion in this area therefore are: what is being done, what can be done, and what should be done to leverage this rapid uptake of mobile and social media technologies among Indigenous Australians for health promotion purposes? Can mobile or social media technologies be effective for health promotion in such communities and, if so, then how? Questions such as these shaped the focus of this initial scoping research.

Methods

Overview

In consultation with a reference group drawn from the Australian Primary Health Care Research Institute (APHCRI) and the Aboriginal Health Council of Western Australia (AHCWA), the research reported in this study aimed to answer the following two research questions: (1) Literature: What is the peer-reviewed evidence of benefit for social media and mobile technologies used in health promotion, intervention, self-management, and health service delivery, with regard to smoking cessation, sexual health, and otitis media? and (2) Projects: What social media and mobile software have been used in Indigenous-focused health promotion interventions in Australia with respect to smoking cessation, sexual health, or otitis media, and what is the evidence of their effectiveness and benefit?

Peer-Reviewed Literature

The scoping study contained two review components. The first was a review of recent publications in peer-reviewed literature, following the framework outlined in Arksey (2003), with the intent to summarize and disseminate research findings and identify gaps in the existing literature. We included both original studies and systematic reviews, looking for evidence of benefit associated with relevant eHealth interventions. The literature search was conducted in November 2013 using Ovid/MEDLINE, searching for publications since 2011 (inclusive) using the following search strategy:

1. exp Health Promotion/ or Health Communication/ or Health Education/ or Social Marketing/ or Health Literacy/ or Preventive Medicine/ or Preventive Health Services/ or Primary Prevention/ or Delivery of Health Care/ or Health Services Accessibility/ or Sexual Behavior/ or Smoking Cessation/ or Smoking/ or Otitis media/
2. exp Social Media/ or Cellular Phone/ or Text Messaging/ or computers, handheld/
3. (facebook or social media or social networking or SMS or text messag$ or smart phone or smartphone or iPhone or iPad or new media or "Web 2.0").af.
4. (facebook or social media or social networking or SMS or text messag$ or smart phone or smartphone or iPhone or iPad or new media or "Web 2.0").m_titl.
5. 2 or 3 or 4
6. 1 and 5

Both primary studies and systematic reviews were collected and assessed by title and abstract for further review.

XSL•FO
RenderX

To be included, studies or reviews had to report on the impact of a health promotion intervention on health or behavioral outcomes, for example: smoking status, sexual health behavior, or improvements in relevant knowledge, attitudes, or intent. Qualitative and quantitative studies were included. All target populations (Indigenous or otherwise) were included.

Included studies used social media (on any platform) or mobile phone software or features (excluding voice features). Automated SMS (short message service) text messaging-only interventions were included on the basis that the basic character of such interventions could reasonably be reproduced using mobile software; other mobile, non-smartphone interventions were included on a similar basis of broad similarity.

Studies of direct communication methods between health services and patients or technological extensions of normal health service activities (eg, reminders or communication of test results) were excluded. Website-delivered interventions (with no social media component) were excluded. Speculative articles and publications not reporting empirical results were excluded. Study protocols and feasibility studies reporting only on acceptability of interventions were also excluded.

Systematic reviews were examined for scope of search, summative evidence reported, and conclusions regarding evidence according to the criteria of this scoping study (outlined above). Cited articles in systematic reviews were assessed for secondary inclusion in the same manner as primary studies.

Data were extracted into a template that recorded the study focus (technological modality and disease focus), method and study population, and evidence of impact or outcome. Multiple publications from the same study were grouped together. Evidence of impact or outcome was broken down according to four broad domains of improvement: knowledge or awareness of relevant health information, health-related attitudes (eg, self-reported behavioral intent or measures of self-efficacy), health-related behaviors, and health outcomes.

Australian Projects

The second component of the scoping exercise looked for (1) publically available mobile apps projects created or promoted by Australian health bodies (including government departments and agencies, Indigenous health organizations, and health promotion agencies) and targeted at Indigenous populations, and (2) social media projects and initiatives created or promoted by Australian health bodies (including government departments and agencies, Indigenous health organizations, and health promotion agencies) and targeted at Indigenous populations.

The primary method used was domain-restricted searches of the websites of these organizations, using the Google search engine with the search terms: “iphone”, “android”, “social media”, and “Facebook”. The term “site:[domain]” was included to search only indexed pages within the domains in question. Descriptions of Indigenous public health projects maintained at the Centre for Excellence in Indigenous Tobacco Control [8] and the Australian Indigenous Health Infonet [9] were manually searched for mentions of social media and mobile software app components. Android and iOS app stores were searched using the search terms “Indigenous”, “Aboriginal”, “Smoking”, and “Health”.

Apps designed to be used by clinicians or other health professionals (eg, medical reference guides or clinical advice support apps) were excluded from the study. Inclusion criteria for both mobile software and social media apps were broad. Any app that appeared to be aimed at an Indigenous audience or which reported intentions in that regard was included. Any use of social media or mobile software apps within larger health promotion projects was included.

In all included cases, we also looked for publically available evaluations of these projects and evidence of benefit, as well as rudimentary measures of reach and uptake including number of “likes” or downloads on the relevant social media platforms and app stores. Process evaluations were not included.

Results

Peer-Reviewed Literature

Our search returned 560 results, which were screened by title and abstract, of which 28 were selected for further examination. Of these, 10 primary intervention studies met our inclusion criteria: six for smoking cessation, four for sexual health, and none for otitis media. Also included were seven systematic reviews: four of a general disease focus, which contained evidence relevant to our inclusion criteria, two specific to smoking cessation, and one specific to sexual health. From these review publications, we secondarily included a further eight publications specific to smoking (drawn from six discrete studies/data sets) and two specific to sexual health.

Two of our nine primary inclusions were also present in review publications; there was one secondary inclusion published during the interval for primary inclusion but not found by our search strategy.

Systematic reviews of a general focus are summarized in Table 1, with all smoking-specific studies in Table 2, and all sexual health-specific studies in Table 3. We found no studies focusing on otitis media that met the inclusion criteria.

One smoking cessation study and two sexual health intervention studies demonstrated some evidence of benefit with both knowledge of relevant health information and health-related attitudes, however none of these three demonstrated improvement in health-related behaviors. Statistically significant improvement in health-related behaviors was reported in four smoking cessation studies. Evidence for behavioral benefit after sexual health interventions was small or tentative and no studies attempted to measure benefits for direct health outcomes.

Australian Projects

We found five projects with significant social media health components targeted at the Indigenous Australian population for health promotion purposes. Of these, four primarily targeted smoking cessation and two targeted sexual health and/or behavior. We found no publically available evaluations of these projects that included evidence beyond that for process evaluation. Available metrics of reach and impact included

XSL•FO
RenderX

Facebook page “likes” ranging between 383 and 2694 users, and 17,143 views in the case of one YouTube video.

We also found 29 examples of mobile phone apps produced, funded, or promoted by Australian health promotion bodies, four of which were identifiably for health promotion purposes targeting the Indigenous population, including three concerned with smoking cessation and behavior and one for otitis media awareness. We found no publically available evaluations of any of these apps. Available metrics of reach and impact from app stores suggested less than 100 installations for the least installed app and more than 5000 for the most popular, which was rated at 4.1/5 and 4.5/5 on the two most popular app stores. There were too few reviews to draw statistical conclusions in the case of the three apps most closely targeted at an Indigenous audience. These results are detailed in Table 4.

Table 1. Summary of general relevance reviews.

Publication	Study summary	Findings/outcomes
Yeager, V & Menachemi, N, 2011. Text messaging in health care: a systematic review of impact studies. [10]	Systematic review of text messaging use in health care; publications prior to 2010.	N=127 articles identified for consideration, n=61 reviewed with n=24 public health related. Publications began in 2003. 7 secondary inclusions: [11-17]
Vodopivec-Jamsek, V et al, 2012. Mobile phone messaging for preventive health care. Cochrane Database of Systematic Reviews. [18]	Systematic review of SMS[a]/mobile messaging – all primary prevention uses, effect on health status and health behavior outcomes. Carried out June 2009.	N=3937 citations, n=4 included interventions on health behaviors (out of 31 studies reviewed). “Very limited evidence”, high quality evidence for smoking cessation intervention only. 1 secondary inclusion: [12]
Chou, WS et al, 2013. Web 2.0 for Health Promotion: Reviewing the Current Evidence. [3]	Systematic review of social media use in health settings. Reviews and commentaries, descriptive studies, intervention studies.	N=1258 citations, n=34 included intervention studies. Only 10 small and/or pilot studies found where social media was used in interventions. 1 secondary inclusion: [19]
Free, C et al, 2013. The Effectiveness of Mobile-Health Technology-Based Health Behaviour Change or Disease Management Interventions for Health Care Consumers: A Systematic Review. [4]	Systematic review of literature 1990 – September 2010. All uses of mobile Meta-analysis including assessment of bias risk.	N=36,314 citations, n=26 included interventions on health behaviors (n=75 total inclusions). SMS-smoking cessation trials (only) are shown to be effective in the meta-analysis. 5 secondary inclusions: [12,13,20-22]

[a]SMS: short message service

XSL•FO
RenderX

Table 2. Summary of smoking cessation reviews.

Publication	Study summary	Findings/outcomes (K=knowledge, A=attitude, B=behavior, H=health)
Primary studies		
Free, C et al, 2011. Smoking cessation support delivered via mobile phone text messaging (txt2stop): a single-blind, randomised trial. [21]	RCT[a] of text messaging program for smoking cessation. N=5800 smokers wishing to quit, UK, 2007-2009. Abstinence biochemically verified.	B - Positive Significant increase in verified abstinence at 6 months, 10.7% intervention vs 4.9% control. Also included via: [4,23]
Reitzel, L et al, 2011. The efficacy of computer-delivered treatment for smoking cessation [24], also: Wetter et al. (2011) A randomized clinical trial of a palmtop computer-delivered treatment for smoking relapse prevention among women. [25]	RCT of palmtop-delivered intervention. USA, 1999-2003. N=303 adult smokers, randomized into palmtop intervention and regular intervention.	B - Negative No significant differences in outcomes between intervention groups.
Whittaker, R et al, 2011. A theory-based video messaging mobile phone intervention for smoking cessation: randomized controlled trial. [26]	RCT of SMS[b] and mobile video content (complex) for smoking cessation. New Zealand, 2007 & 2009. N=226 adults recruited via advertising.	B - Negative Not able to demonstrate statistically significant effect for complex/tailored message and video regime over control. Failed to reach sample size. Also included via: [23]
Naughton, F et al, 2012. Randomized controlled trial evaluation of a tailored leaflet and SMS text message self-help intervention for pregnant smokers (MiQuit). [27]	RCT of text messaging as part of support package for smoking cessation. UK, 2008-2009. N=207 pregnant smokers. Intervention of tailored literature and intervention SMS vs standard self-help literature and evaluation SMS.	K - Positive, A - Positive, B - Negative Positive effects on "key intervention targets: self-efficacy, harm beliefs, and determination to quit", and increased likelihood of setting quit date. Self-reported and biochemically verified abstinence rate difference not statistically significant. Study was "not powered to detect a group difference on smoking outcomes".
Ybarra, M et al, 2012. A text messaging-based smoking cessation program for adult smokers: randomized controlled trial. [28]	RCT of text messaging for smoking cessation. Turkey. N=151 adult smokers wanting to quit. Randomized into text messaging and brochure groups, 3-month follow-up.	B - Mostly negative Overall biochemically verified 3-month cessation trends not significantly higher in intervention group, except for women where 5 female participants quit (15%), significantly more than in control (0).
Systematic reviews		
Chen, Y et al, 2012. Effectiveness and cost-effectiveness of computer and other electronic aids for smoking cessation: a systematic review and network meta-analysis. [29]	Broad systematic review and network meta-analysis of all "computer and electronic aids" for smoking cessation. Inclusions up to 2009.	B - Positive/tentative N=3969 citations, 77 papers from 60 RCTs included. Effectiveness review: studied aids increase likelihood of cessation "but the effect is small". Likely to be cost effective for smokers seeking to quit. Not able to determine what form of electronic aids are more effective. 5 secondary inclusions: [12-15,20]
Whittaker, R et al, 2012. Mobile phone-based interventions for smoking cessation. [23]	Narrow Cochrane review of long term (6 month+) studies of text messaging interventions for smoking cessation, carried out May 2012.	B - Positive/mixed 5 studies included out of 17 reviewed. Three used the same "text2quit" SMS intervention adapted for different populations and contexts (Rogers 2005; Free 2009; Free 2011). 4 secondary inclusions: [12,20,21,26]
Secondary inclusions from systematic reviews		
Obermayer, JL et al, 2004. College Smoking-Cessation Using Cell Phone Text Messaging. [16]	Cohort study, N=46 young adult smokers, Web and text messaging "quit program" intervention for cessation.	B: Tentative Small, uncontrolled pilot study with modest outcome measures. 43% participants attempted to quit for at least 24h and 17% had been abstinent for 1 week after 6 weeks. Included via [10].

XSL•FO
RenderX

Publication	Study summary	Findings/outcomes (K=knowledge, A=attitude, B=behavior, H=health)
Rodgers et al (2005) Do u smoke after txt? Results of a randomised trial of smoking cessation using mobile phone text messaging [12], and: Bramley et al (2005) Smoking cessation using mobile phone text messaging is as effective in Maori as non-Maori. [11]	RCT, N=1705 adult New Zealanders, 26-week text messaging intervention for cessation.	B: Positive More participants had quit at 6 weeks in the intervention compared to control group: 239 (28%) vs 109 (13%). Intervention as effective in Maori as non-Maori. Long-term results compromised by incomplete follow-up and over-reporting of quit rate. Included via [4,10,18,23,29].
Brendryen and Kraft (2008) Happy Ending: a randomized controlled trial of a digital multi-media smoking cessation intervention [14], and: Brendryen et al (2008) A Digital Smoking Cessation Program Delivered Through Internet and Cell Phone Without Nicotine Replacement (Happy Ending): Randomized Controlled Trial. [15]	RCT, Norway. N=396 smokers, 56 week automated text message, email, and Web-based smoking cessation intervention.	B: Positive Statistically significantly higher repeated point abstinence rates (at 1, 3, 6 and 12 months) than control participants (22.3% vs 13.1%). Included via [10,29].
Whittaker, R et al, 2008. A Multimedia Mobile Phone-Based Youth Smoking Cessation Intervention: Findings From Content Development and Piloting Studies. [17]	Feasibility study and uncontrolled pilot study. New Zealand. N=15 young smokers, 4-week video messaging program for smoking cessation.	B - Tentative Mostly a feasibility study. Positive feedback from pilot group suggesting improved smoking abstinence (not verified). Included via [10].
Haug, S et al., 2009. Continuous individual support of smoking cessation using text messaging: A pilot experimental study. [13]	RCT of text messaging for smoking cessation. Germany. N=174 adults, weekly text message question followed by 1 or 3 feedback text messages and helpline.	B - Negative No significant differences found in smoking variables around abstinence. Primarily a feasibility study. Included via [4,10,29].
Free, C et al., 2009. Txt2stop: a pilot randomised controlled trial of mobile phone-based smoking cessation support. [20]	Pilot of RCT of text messaging program for smoking cessation [21]. UK. N=200 smokers wishing to quit. Abstinence biochemically verified.	B - Positive Abstinence rates at 4 weeks significantly better for intervention group: 26% vs 12%. Included via [4,23,29].

[a]RCT: randomized controlled trial

[b]SMS: short message service

XSL•FO
RenderX

Table 3. Summary of sexual health reviews.

Publication	Study Summary	Findings/outcomes (K=knowledge, A=attitude, B=behavior, H=health)
Primary inclusions		
Juzang, I et al, 2011. A pilot programme using mobile phones for HIV prevention. [30]	Pilot/feasibility trial of text messaging for sexual health knowledge and behavior. USA. N=60 young black men. 12 week trial, intervention received sex-ed messages, control received nutrition messages.	K - Positive, A - Positive, B - Negative Participants showed trends in increased monogamy at follow-up compared to controls. Awareness of sexual health significantly higher in the intervention group. Condom norms significantly higher for control. No differences in proportion of protected sex acts.
Gold, J et al, 2011. A randomised controlled trial using mobile advertising to promote safer sex and sun safety to young people. [31]	RCT[a]of SMS[b]advertising for sexual health behavior. Victoria, Australia, 2009. N=7606 16-29 y/o mobile advertising subscribers randomized into “sex” or “sun” groups. Baseline and follow-up mobile phone questionnaire.	B - Mixed Significant OR (self-reported) for less sexually risky behavior (eg fewer sexual partners, always using condoms). Increased STI testing not found. Only 10% of participants completed any surveys, only 151 completed both baseline and follow-up surveys. Authors cite “challenges experienced during intervention implementation”.
Bull, S et al, 2012. Social media-delivered sexual health intervention: a cluster randomized controlled trial. [32]	Cluster RCT trial of Facebook messages for condom use. USA, October 2010-May 2011. N=1578 self/peer-recruited enrollees to a Facebook page with sexual health messages. Condom use surveyed at 2 and 6 months.	A - Negative, B - Positive/Tentative Small differences in condom use seen at 2 months (eg, 63% vs 57%, *P*=.03), no effects seen at 6 months. No effect on self-efficacy or norms.
Reback, C et al, 2012. Text messaging reduces HIV risk behaviors among methamphetamine-using men who have sex with men. [33]	Non-controlled study of text messages targeting risky sexual behavior and drug use. USA October 2008-May 2009. N=52 methamphetamine-using men who have sex with men. Text messages over 2 weeks.	B - Tentative Self-report behavior significantly improved at 2 months. Biochemical verification of self-reported outcomes not statistically significant. Not controlled, baseline follow up only.
Systematic reviews		
Guse, K et al, 2012. Interventions using new digital media to improve adolescent sexual health: a systematic review. [34]	Broad review of interactive digital media (new media) for adolescent sexual health.	B - Positive/mixed N=942 citations, 10 included. Mostly Web-delivered interactive interventions, three showing significant impact in dangerous sexual behaviors. 1 secondary inclusion: [19]
Secondary inclusions		
Moreno, MA et al, 2009. Reducing at-risk adolescents’ display of risk behavior on a social networking Web site: a randomized controlled pilot intervention trial. [19]	RCT of intervention to alter online displays of risky behavior. N=190 18-20 year olds with MySpace profiles. Intervention group sent single message from clinician.	B - Tentative/negative Online displays of risky sexual behavior slightly reduced 3 months after intervention, no significant change in other measures. Only online behaviors studied. Included via: [3,34]
Lim, MSC et al, 2012. Impact of text and email messaging on the sexual health of young people: a randomised controlled trial. [22]	RCT on 12-month program of email and text messages to young adults. S&P: N=994 recruited at music festival. Self-reported behavior and knowledge via survey at recruitment, 3, 6, and 12 months.	K - Positive, A - Positive, B - Negative At 12 months, STI knowledge was higher in the intervention. Women (but not men) in the intervention group were more likely to have had an STI test or discuss sexual health with a clinician. No significant impact on condom use. Included via: [4]

[a]RCT: randomized controlled trial

[b]SMS: short message service

XSL•FO
RenderX

Table 4. Australian apps and social media programs with Indigenous focus.

Name of app or campaign	Organization	Description	Evaluation or evidence of reach / impact
Stickin' it up the smokes	Aboriginal Health Council of South Australia	Social Media – Smoking. Social marketing campaign with prominent Facebook page, targeting smoking cessation/abstinence for young Aboriginal women.	Reach/impact evaluation not available at time of study. Facebook page has 1274 likes, 0-19 likes per post.
It's your choice! Have a voice! rights, respect, responsibility 2013	Aboriginal Health and Medical Research Council	Social Media – Sexual Health. Publicity campaign partnered with Indigenous Hip Hop Projects (IHHP), "bringing a state-wide arts based campaign to empower and educate Aboriginal adolescents to make informed choices about sexual and reproductive health and understand the negative impacts of alcohol and other drugs". Facebook page included as part of broader campaign.	No publically available evaluation found. Facebook page has 2694 likes, 0-54 likes per post.
Kasa por yarn	Funded by Queensland Health.	Social Media – Sexual Health. Audio/video drama with YouTube and Facebook presence, funded by Queensland Health. Two series on Torres Strait Radio 4MW in 2011 and 2012. Storylines and characters were locally developed on Thursday Island.	No publically available evaluation found. Facebook page active 2011-2012, currently has 2639 likes. YouTube channel's 42 videos have between 88 and 1936 views.
Rewrite your story	Nunkuwarrin Yunti of South Australia Inc.	Social Media – Smoking. Web–based campaign focused on social media, launched January 2013. Includes sophisticated website with interactive "pledge" feature, Facebook page, and YouTube channel hosting personal "stories" about smoking and smoking cessation.	No publically available evaluation found. Main website includes 402 "pledges and stories". Facebook page active 2011-present, currently has 443 likes. YouTube channel's 20 videos have between 8 and 1793 views.
NoSmokes.com.au	Menzies School of Health Research	Social Media/Mobile Software – Smoking. Suite of online projects /experiments designed for use by Aboriginal and Torres Strait Islander people, including mobile software, videos, and online games. Hosted from dedicated website, Facebook page, and YouTube channel.	Only focus-group/process evaluation currently available. Facebook page active 2010 to present has 383 likes, YouTube channel's 33 videos have between 9 and 17,143 views.
Hip Hop Dance-Off	Menzies School of Health Research	Mobile Software – Smoking. Part of "No Smokes" suite of eHealth projects.	10 ratings on iTunes store, 1000-5000 installs on Google Play store.
No Smokes/So you think you can Quit?	Menzies School of Health Research	Mobile Software – Smoking. Part of "No Smokes" suite of eHealth projects. App available for iPhone, iPad, and Android.	9 ratings on iTunes store, 10-50 installs on Google Play store.
Quit for you, quit for two	Commonwealth Department of Health	Mobile Software - Smoking. Mobile app, part of government advertising campaign intended to encourage mothers from a "diverse background" to quit smoking. Includes tracker/educational component for baby progress and money saved, and an animated baby character will play games, assist with timing breathing, etc. Includes Quitline connection and other support options.	5000-10,000 installs on Google Play store with 21 ratings at 4.1/5, 6 ratings on iTunes store at 4.5/5.

XSL•FO
RenderX

Name of app or campaign	Organization	Description	Evaluation or evidence of reach / impact
Talking Book/Care for Kids Ears	Commonwealth Department of Health	Mobile Software – Otitis Media. Basic ear health information presented in style of an interactive children's book, read in English or many Indigenous languages.	100-200 installs on Google Play store, insufficient ratings on iTunes store.

Discussion

Summary of Findings

With our survey of the relevant peer-reviewed literature, our main finding was that the evidence of benefit for social media and mobile apps in the specific areas of interest for Indigenous health in Australia is, for the most part, tentative and scattered. The most compelling body of evidence, relevant to this study, comes from a series of international studies into text messaging for smoking cessation. The main positive results here begin with a 2005 study in New Zealand [11,12], which was then adapted for further studies in the United Kingdom [20,21] and Turkey [28]. While these results are promising for messaging-based encouragement programs (and apps that might mimic such functionality via mobile software), this is largely an isolated body of research on messaging interventions specifically, and the transferability of such results from the context of communication to purely automated applications is not clear.

In fact, no peer-reviewed research was found at all for mobile software apps as such and only a single study was found where a social media intervention was linked to a significant (though small) change in behaviors that were *directly* health-relevant [32]. The vast majority of included studies that examined evidence of benefit studied the use of SMS text or similar messaging services to provide reminders, education, or preventive health messages. The key technologies of interest with the lowest barriers of entry and highest potential for dissemination—mobile app software and social media—do not yet appear to have a significant body of peer-reviewed evidence examining their effectiveness.

Our scoping study found four extant uses of social media in specific public health projects targeting an Indigenous audience, in most cases utilizing ready-made platforms such as Facebook and YouTube. The four mobile software apps that met the inclusion criteria included two produced by the Australian Commonwealth Government, and two produced as part of a broader research effort into eHealth for smoking cessation: the "No Smokes" project, which is also included in Table 4.

Among these projects we found a mix of strategies, including health promotion content produced with participation from members of the intended audience (eg, Kasa por yarn), and social media use as an addition to a large health campaign touring schools and communities (It's your choice! Have a voice!). In contrast, the projects from "No Smokes" and "Rewrite your story" appeared to rely on a more top-down or Web-based approach to dissemination. The two Commonwealth-funded mobile software projects also differed in how narrowly they focused on the Indigenous population, with "Care for kids' ears" explicitly aimed at Indigenous children and caregivers, but the more widely used "Quit for you, quit for two" app appealing to a much broader audience and being only loosely tied to the government's target to reduce the Indigenous smoking rate.

The Limited Scope of Evidence

Of the text messaging studies, a significant number have shown robust results for interventions to aid in smoking cessation for smokers trying to quit, either on their own or as part of a broader "quit smoking" program—a result confirmed in a recent systematic review and meta-analysis [4]. There is less evidence however for text messaging interventions for sexual health and other areas, though some studies here are also promising.

However, this imbalance of the evidence in favor of text messaging interventions is no doubt partly due to the relative maturity and ubiquity of this mobile technology (compared to mobile app software and social media), and may also speak to specific difficulties with conducting research on the newer technologies and/or publishing such research in peer-reviewed journals.

In any case, the lack of peer-reviewed or publically available evidence for interactive social media and mobile software projects is especially curious given the number of such projects being implemented in Australia alone.

We suggest that there are three distinct issues that may help explain this lack of evidence.

Study Limitations and the Terminology Problem

One distinguishing feature of research in this area is that the language itself can be fragmented and ambiguous. For example, the term "eHealth" is one of the oldest and most well-established, but ambiguously refers to many different kinds of communications technology: technologies for connecting consumers and health services, for interconnecting health services, for training purposes, or simply as an extension of eCommerce (ie, as advertising/marketing/sales systems for commercial health services [35]).

This ambiguity within the research literature alone has itself been an object of study and been recognized in systematic reviews, as summarized in a 2005 review by Oh et al: "As with most neologisms, the precise meaning of eHealth varied with the context in which the term was used. Nevertheless, it has been fairly well understood, and is now widely used by many academic institutions, professional bodies, and funding organizations. We recognized the impossibility of finding an universally acceptable, universally applicable formal definition, yet felt that a clearer understanding of the term could be achieved by reviewing the range of proposed meanings" [36].

XSL•FO
RenderX

The result of this phenomenon is that multiple definitions and formalizations have been used by researchers to try to capture only partially overlapping concepts, with little or no systematic approach between those researchers and consumers, health services, and content creators. This has also seen new terms being coined both in the health context (eg, mHealth for mobile phone-based technology use) and in the broader community to describe new trends in the overall technology marketplace (eg, Web 2.0). So the terms researchers might seek to use and their referential extension over time can be extremely volatile. Furthermore, the rapid growth and aggressive use of branding in this area of commerce and technology has created what might be called the Hoover, Xerox, or iPod effect: where brand names and proprietary eponyms like “Facebook” dominate (and in some cases predate) the use of generic terms like “social media”. In many of the titles and abstracts considered in this study, the brand names of technologies used (eg, Facebook) were the only text-searchable identifier of relevance, suggesting that standardized nomenclature in research is struggling to catch up with its subject matter.

It is beyond the scope of this study to consider whether this terminologically heterogeneous research environment implies genuine conceptual fragmentation, however, at a practical level at least, it certainly means that search strategies must either focus narrowly or aim for being thorough at the expense of subsequent workload. There are no good MeSH terms that capture the area of interest presented in this review and no simple, stable language for keyword text searches. For example, one recent systematic review used a search strategy with keywords that took up 2 pages and returned over 26,000 electronic citations for review, out of which 75 were eventually deemed relevant [4]. In terms of workload alone, this represents a formidable challenge for researchers.

The Measures Problem

A second issue is that the nature of the technology being used does not allow for clear implementation measures or clear outcome measures. For example, social media, websites, and app software are “use on demand” services in that they are engaged with by users to varying degrees and under a variety of conditions, so there is no “standard dosage” when seen as interventions. Likewise, these technologies provide a degree of distance and anonymity for subjects both inimical to following up outcomes (we do not know who is using it “in the wild”) and difficult to reproduce in controlled environments (normal conditions of use are very unlike controlled conditions).

These issues may help explain the lack of relatively robust evidence for such interventions when compared to text messaging interventions where dosage (ie, the messages sent and their content) is largely under the control of the researchers and can be standardized (with some variation) across hundreds of participants. We would argue that the very nature of interactive media here undermines the scope for large scale intervention studies.

The Institutional Problem

More speculatively, we suggest this will imply significant *institutional* hurdles that can be expected to work against the successful publication of research in this area, thereby lowering incentives for research investment.

The most obvious impediment is with regard to what is taken to be good evidence in health care research in general, when combined with the measures problem as described above. Because of the measures problem, it is difficult to conduct a randomized controlled trial of a social media or mobile app intervention that simultaneously approaches the gold standard for evidence in the health sciences and also studies real-world engagement with the intervention. Likewise, because real-world dissemination of apps and social media is user-driven, the capacity for robust recruitment and follow-up methodology is greatly impaired.

This is in stark contrast to text messaging interventions where dosage can be firmly under the control of researchers and delivered directly to recruited test subjects who can be directly contacted for follow-up purposes. The same features of social media and mobile software that give them great ease of use and massive disseminability therefore also make them relatively inimitable to traditional academic research. Social media and mobile software use is too slippery a fish for the standard research nets to catch.

So we can expect there to be a systematic barrier against the production of peer-reviewed publications in this area, which may be acting to de-incentivize much-needed research. This is not to say that evidence itself in this field is not possible and program evaluations are generally expected to be built in to publically funded health promotion projects (ie, to gather evidence regardless of foreseen academic output). Measures of reach and usage for apps and social media can be easily obtained from social media analytics and well-designed software to at least gain an understanding of the degree of uptake. However, even if good “reach” data is obtained, there is still a problem of recruitment for the sake of comparative data: recruiting the app/social media users for assessment of health, knowledge, or behavior status. And randomization into intervention and control groups faces further significant technical hurdles. These are fundamental challenges for both research and evaluation in this area, though there are some encouraging signs that innovative researchers can find ways to approach them in a robust manner [32].

One final piece of evidence in this regard concerns timeframe and turnaround of the research itself. Many of the studies we found were published several years after the initial data collection had been carried out and the systematic reviews found were similarly published at a considerable remove from the most recent publications reviewed in each. The optimistic outlook here is that perhaps this slow process is concealing a coming wave of recent mobile software or social media studies. However, it also further diminishes the value of conducting research in such a rapidly changing field; if not carefully designed for general applicability, by the time the research is published there is a non-trivial chance that the studied intervention is obsolete or that the target audience has moved on.

XSL•FO
RenderX

Future Research

More research on social media and mobile software interventions is required to establish how effective these interventions are for health promotion purposes and under what conditions they can be made more so. However, there is also reason to think that there are several specific impediments to research in this area, which themselves call for further research.

The task for evidence-seeking research and evaluation projects in consumer-focused eHealth intervention is a challenging one, consisting of three distinct requirements: (1) incorporate all useful measures of reach, behavior, and use, (2) make a connection between these measures and (some measure of) health outcomes or health-relevant behavior, and (3) conduct this research in a manner that is comprehensible and rapidly disseminable for peer review.

At the abstract level, we have suggested that there are three areas that will need to be addressed. The process of doing this may require deeper health researcher engagement with the relevant literature in media studies, information technology, and eCommerce, to clarify the concepts and methods involved, converge more upon a common technical language, and develop evidential standards for the use of software analytics and other evidence not commonly found in health and medical research. This might be seen as a need for relevant research in knowledge translation and implementation: to better understand how to translate the principles of commercial success in social media and mobile software into effective health promotion interventions, and how to better integrate these methods into health research.

At a more applied level, there is greater urgency for researchers and health promotion professionals to cooperate in the evaluation of social media and mobile software interventions. As we found, there are numerous projects in Australia alone that could constitute valuable experiments if the right data were also able to be gathered and the results made applicable for peer review and public dissemination.

Concerning the limitations of this scoping study, while we are satisfied that we have sampled a suitable cross-section of research literature and existing projects for our particular purposes, our focus on specific health topics make it likely we have missed some relevant studies and initiatives. Future studies into similar literature (eg, investigating different target conditions) would be advised to carefully consider search strategies as used in the many systematic reviews in eHealth now available. Our survey of Indigenous health programs was also limited to examining publically available information and will not have captured evaluation activity that is underway or planned.

Indigenous Australia

Finally, we conclude by returning to consider our original framing question: how might social media and mobile technologies best be leveraged for maximum reach and best health outcomes among Indigenous populations? To answer this question would be to consider one further issue inherent to the interactive and collaborative nature of social media and mobile technologies: the unpredictable and culturally-specific nature of their use.

The adoption of an online social media platform is a paradigmatically chaotic social process and appears to occur when one becomes the dominant available space for an already socially self-identifying population to move (at least some of) their social networking activities into. It is well understood that the uptake of social media platforms and mobile software has varied widely according to language, culture, and demographics, and for a variety of sometimes surprising reasons. Usage is driven by the utility of social connections and coherence, but unexpected contingencies often seem to be deciding factors, which explains, for example, why Facebook dominates in English-speaking countries, while it was Orkut (an early attempt at a social network by Google) that rose to early dominance in Brazil and India [37]. Among some Indigenous communities in the Northern Territory, a similar contingency occurred with the rapid uptake of “Divas Chat”, a minimal messaging and social network platform attached to Telstra pre-paid services, but which offers greater reliability in such remote communities [38,39].

Observations such as this (and consideration of the cultural uniqueness of rural and remote Indigenous communities) raise several questions relevant to researchers looking at this area. Of particular interest here is: do culturally-specific modes and organizing principles around “offline” social networking produce specific ways of engaging with online social media and/or absorbing the content therein?

Addressing considerations such as these is certainly not just the domain of health and medical Internet researchers; they touch on how the introduction of modern media influences a community and how technology coexists with culture. Therefore, a full understanding of what might work, and how, may require cross-disciplinary research including the contributions from, for example, cultural researchers and social anthropologists—and, of course, the communities themselves.

Conclusions

Social media and mobile software interventions are already being used for health promotion and appear to hold great promise, especially for Indigenous or other traditionally underserved populations. However, the current evidence for their effectiveness or health benefit is sparse and mixed. This lack of evidence should not necessarily be seen as an indictment, as it is perhaps to be expected given the mercurial and elusive nature of these interventions and how target populations engage with them. Nevertheless, intervention projects being developed in this area, no matter how well thought out or enthusiastically pursued, cannot be described as entirely evidence-based given the current state of the evidence.

However, for health promotion, there is no real alternative but to engage with social media and mobile software technology, as these forms of online interaction are becoming increasingly ubiquitous at the expense of more traditional media. This means that a more thorough and professional understanding of these technologies will increasingly be called for, at the level of the

XSL•FO
RenderX

technologies themselves, but also with regard to how they are in turn engaged with by the specific populations of interest.

Acknowledgments

This study was conducted as part of a research partnership between the Aboriginal Health Council of Western Australia and the Australian Primary Health Care Research Institute at the Australian National University. We wish to acknowledge Dr Beverley Sibthorpe and Ms Donisha Duff who provided advice to this project through our reference group and Ms Mier Chan who provided administrative support and assistance.

Conflicts of Interest

None declared.

References

1. Grier S, Bryant CA. Social marketing in public health. Annu Rev Public Health 2005;26:319-339. [doi: 10.1146/annurev.publhealth.26.021304.144610] [Medline: 15760292]
2. Suarez-Almazor ME. Changing health behaviors with social marketing. Osteoporos Int 2011 Aug;22 Suppl 3:461-463. [doi: 10.1007/s00198-011-1699-6] [Medline: 21847766]
3. Chou WY, Prestin A, Lyons C, Wen KY. Web 2.0 for health promotion: reviewing the current evidence. Am J Public Health 2013 Jan;103(1):e9-18. [doi: 10.2105/AJPH.2012.301071] [Medline: 23153164]
4. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. PLoS Med 2013;10(1):e1001362 [FREE Full text] [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]
5. Gartner Newsroom. 2013 Oct 21. Gartner says worldwide PC, tablet and mobile phone shipments to grow 4.5 percent in 2013 as lower-priced devices drive growth URL: http://www.gartner.com/newsroom/id/2610015 [accessed 2014-05-25] [WebCite Cache ID 6PpgFCqvu]
6. Johnson G. Technology use and reading comprehension among Australian Indigenous adolescents. International Journal of Economy, Management and Social Sciences 2013;2(8):558-564 [FREE Full text]
7. Healy JDL. Yolngu Zorba meets Superman. Anthrovision Vaneasa Online Journal 2013;1(1):1-18 [FREE Full text]
8. Centre for Excellence in Indigenous Tobacco Control. 2014 May 25. Indigenous Projects Register URL: http://www.ceitc.org.au/indigenous-projects-register [accessed 2014-05-25] [WebCite Cache ID 6PpculPMa]
9. Australian Indigenous HealthInfoNet. 2014 May 25. Programs and projects - Key resources URL: http://www.healthinfonet.ecu.edu.au/key-resources/programs-projects [accessed 2014-05-25] [WebCite Cache ID 6PpdKalIh]
10. Yeager VA, Menachemi N. Text messaging in health care: a systematic review of impact studies. Adv Health Care Manag 2011;11:235-261. [Medline: 22908672]
11. Bramley D, Riddell T, Whittaker R, Corbett T, Lin RB, Wills M, et al. Smoking cessation using mobile phone text messaging is as effective in Maori as non-Maori. N Z Med J 2005 Jun 3;118(1216):U1494. [Medline: 15937529]
12. Rodgers A, Corbett T, Bramley D, Riddell T, Wills M, Lin RB, 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]
13. Haug S, Meyer C, Schorr G, Bauer S, John U. Continuous individual support of smoking cessation using text messaging: a pilot experimental study. Nicotine Tob Res 2009 Aug;11(8):915-923. [doi: 10.1093/ntr/ntp084] [Medline: 19542517]
14. 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]
15. 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]
16. Obermayer JL, Riley WT, Asif O, Jean-Mary J. College smoking-cessation using cell phone text messaging. J Am Coll Health 2004;53(2):71-78. [doi: 10.3200/JACH.53.2.71-78] [Medline: 15495883]
17. Whittaker R, Maddison R, McRobbie H, Bullen C, Denny S, Dorey E, et al. A multimedia mobile phone-based youth smoking cessation intervention: findings from content development and piloting studies. J Med Internet Res 2008;10(5):e49 [FREE Full text] [doi: 10.2196/jmir.1007] [Medline: 19033148]
18. Vodopivec-Jamsek V, de Jongh T, Gurol-Urganci I, Atun R, Car J. Mobile phone messaging for preventive health care. Cochrane Database Syst Rev 2012;12:CD007457. [doi: 10.1002/14651858.CD007457.pub2] [Medline: 23235643]
19. Moreno MA, Vanderstoep A, Parks MR, Zimmerman FJ, Kurth A, Christakis DA. Reducing at-risk adolescents' display of risk behavior on a social networking web site: a randomized controlled pilot intervention trial. Arch Pediatr Adolesc Med 2009 Jan;163(1):35-41. [doi: 10.1001/archpediatrics.2008.502] [Medline: 19124701]

XSL•FO
RenderX

20. 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]
21. Free C, Knight R, Robertson S, Whittaker R, Edwards P, Zhou W, et al. Smoking cessation support delivered via mobile phone text messaging (txt2stop): a single-blind, randomised trial. Lancet 2011 Jul 2;378(9785):49-55 [FREE Full text] [doi: 10.1016/S0140-6736(11)60701-0] [Medline: 21722952]
22. Lim MS, Hocking JS, Aitken CK, Fairley CK, Jordan L, Lewis JA, et al. Impact of text and email messaging on the sexual health of young people: a randomised controlled trial. J Epidemiol Community Health 2012 Jan;66(1):69-74. [doi: 10.1136/jech.2009.100396] [Medline: 21415232]
23. Whittaker R, Borland R, Bullen C, Lin R, McRobbie H, Rodgers A. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2009(4):CD006611. [doi: 10.1002/14651858.CD006611.pub2] [Medline: 19821377]
24. Reitzel LR, McClure JB, Cofta-Woerpel L, Mazas CA, Cao Y, Cinciripini PM, et al. The efficacy of computer-delivered treatment for smoking cessation. Cancer Epidemiol Biomarkers Prev 2011 Jul;20(7):1555-1557 [FREE Full text] [doi: 10.1158/1055-9965.EPI-11-0390] [Medline: 21613389]
25. Wetter DW, McClure JB, Cofta-Woerpel L, Costello TJ, Reitzel LR, Businelle MS, et al. A randomized clinical trial of a palmtop computer-delivered treatment for smoking relapse prevention among women. Psychol Addict Behav 2011 Jun;25(2):365-371. [doi: 10.1037/a0022797]
26. Whittaker R, Dorey E, Bramley D, Bullen C, Denny S, Elley CR, et al. A theory-based video messaging mobile phone intervention for smoking cessation: randomized controlled trial. J Med Internet Res 2011;13(1):e10 [FREE Full text] [doi: 10.2196/jmir.1553] [Medline: 21371991]
27. Naughton F, Prevost AT, Gilbert H, Sutton S. Randomized controlled trial evaluation of a tailored leaflet and SMS text message self-help intervention for pregnant smokers (MiQuit). Nicotine Tob Res 2012 May;14(5):569-577. [doi: 10.1093/ntr/ntr254] [Medline: 22311960]
28. Ybarra M, Bağci Bosi AT, Korchmaros J, Emri S. A text messaging-based smoking cessation program for adult smokers: randomized controlled trial. J Med Internet Res 2012;14(6):e172 [FREE Full text] [doi: 10.2196/jmir.2231] [Medline: 23271159]
29. Chen YF, Madan J, Welton N, Yahaya I, Aveyard P, Bauld L, et al. Effectiveness and cost-effectiveness of computer and other electronic aids for smoking cessation: a systematic review and network meta-analysis. Health Technol Assess 2012;16(38):1-205, iii [FREE Full text] [doi: 10.3310/hta16380] [Medline: 23046909]
30. Juzang I, Fortune T, Black S, Wright E, Bull S. A pilot programme using mobile phones for HIV prevention. J Telemed Telecare 2011;17(3):150-153. [doi: 10.1258/jtt.2010.091107] [Medline: 21270049]
31. Gold J, Aitken CK, Dixon HG, Lim MS, Gouillou M, Spelman T, et al. A randomised controlled trial using mobile advertising to promote safer sex and sun safety to young people. Health Educ Res 2011 Oct;26(5):782-794 [FREE Full text] [doi: 10.1093/her/cyr020] [Medline: 21447750]
32. Bull SS, Levine DK, Black SR, Schmiege SJ, Santelli J. Social media-delivered sexual health intervention: a cluster randomized controlled trial. Am J Prev Med 2012 Nov;43(5):467-474 [FREE Full text] [doi: 10.1016/j.amepre.2012.07.022] [Medline: 23079168]
33. Reback CJ, Grant DL, Fletcher JB, Branson CM, Shoptaw S, Bowers JR, et al. Text messaging reduces HIV risk behaviors among methamphetamine-using men who have sex with men. AIDS Behav 2012 Oct;16(7):1993-2002. [doi: 10.1007/s10461-012-0200-7] [Medline: 22610370]
34. Guse K, Levine D, Martins S, Lira A, Gaarde J, Westmorland W, et al. Interventions using new digital media to improve adolescent sexual health: a systematic review. J Adolesc Health 2012 Dec;51(6):535-543. [doi: 10.1016/j.jadohealth.2012.03.014] [Medline: 23174462]
35. Oh H, Rizo C, Enkin M, Jadad A. What is eHealth (3): a systematic review of published definitions. J Med Internet Res 2005;7(1):e1 [FREE Full text] [doi: 10.2196/jmir.7.1.e1] [Medline: 15829471]
36. Jones R, Rogers R, Roberts J, Callaghan L, Lindsey L, Campbell J, et al. What is eHealth (5): a research agenda for eHealth through stakeholder consultation and policy context review. J Med Internet Res 2005;7(5):e54 [FREE Full text] [doi: 10.2196/jmir.7.5.e54] [Medline: 16403718]
37. Cosenza V. Vincos Blog. 2014 Jan 20. World map of social networks URL: http://vincos.it/world-map-of-social-networks/ [accessed 2014-05-25] [WebCite Cache ID 6PpmgPZc6]
38. Rawlinson C. ABC Darwin - Australian Broadcasting Corporation. 2013 Dec 15. "Divas Chat" re-shaping remote communities URL: http://www.abc.net.au/local/stories/2011/12/22/3396787.htm [accessed 2014-05-25] [WebCite Cache ID 6PpmoZ02v]
39. Quartermaine C. SBS News. 2013. Divas Chat App defrauding women in remote communities URL: http://www.sbs.com.au/news/article/2013/11/25/divas-chat-app-defrauding-women-remote-communities [accessed 2014-05-25] [WebCite Cache ID 6Ppn2psT1]

Abbreviations

RCT: randomized controlled trial

XSL•FO
RenderX

SMS: short message service

Edited by G Eysenbach; submitted 19.06.14; peer-reviewed by B Carlson, J McCalman; comments to author 24.09.14; revised version received 13.10.14; accepted 28.10.14; published 10.12.14.

Please cite as:
Brusse C, Gardner K, McAullay D, Dowden M
Social Media and Mobile Apps for Health Promotion in Australian Indigenous Populations: Scoping Review
J Med Internet Res 2014;16(12):e280
URL: http://www.jmir.org/2014/12/e280/
doi:10.2196/jmir.3614
PMID:25498835

©Carl Brusse, Karen Gardner, Daniel McAullay, Michelle Dowden. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 10.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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 Effects in Regular Practice of E-Communication and Web-Based Self-Management Support Among Breast Cancer Patients: Preliminary Results From a Randomized Controlled Trial

Elin Børøsund[1,2], MSc; Milada Cvancarova[1], PhD; Shirley M Moore[3], PhD; Mirjam Ekstedt[1,4], PhD; Cornelia M Ruland[1,2], PhD

[1]Centre for Shared Decision Making and Collaborative Care Research, Division of Medicine, Oslo University Hospital, Oslo, Norway

[2]Department of Medicine, University of Oslo, Oslo, Norway

[3]Frances Payne Bolton School of Nursing, Case Western Reserve University, Cleveland, OH, United States

[4]Royal Institute of Technology, KTH, School of Technology and Health, Stockholm, Sweden

Corresponding Author:
Cornelia M Ruland, PhD
Centre for Shared Decision Making and Collaborative Care Research
Division of Medicine
Oslo University Hospital
Pb 4950 Nydalen
Oslo, 0424
Norway
Phone: 47 23075460
Fax: 47 23075450
Email: cornelia.ruland@rr-research.no

Abstract

Background: While Web-based interventions have been shown to assist a wide range of patients successfully in managing their illness, few studies have examined the relative contribution of different Web-based components to improve outcomes. Further efficacy trials are needed to test the effects of Web support when offered as a part of routine care.

Objective: Our aim was to compare in regular care the effects of (1) an Internet-based patient provider communication service (IPPC), (2) WebChoice, a Web-based illness management system for breast cancer patients (IPPC included), and (3) usual care on symptom distress, anxiety, depression, (primary outcomes), and self-efficacy (secondary outcome). This study reports preliminary findings from 6 months' follow-up data in a 12-month trial.

Methods: We recruited 167 patients recently diagnosed with breast cancer and undergoing treatment from three Norwegian hospitals. The nurse-administered IPPC allowed patients to send secure e-messages to and receive e-messages from health care personnel at the hospital where they were treated. In addition to the IPPC, WebChoice contains components for symptom monitoring, tailored information and self-management support, a diary, and communication with other patients. A total of 20 care providers (11 nurses, 6 physicians, and 3 social workers) were trained to answer questions from patients. Outcomes were measured with questionnaires at study entry and at study months 2, 4, and 6. Linear mixed models for repeated measures were fitted to compare effects on outcomes over time.

Results: Patients were randomly assigned to the WebChoice group (n=64), the IPPC group (n=45), or the usual care group (n=58). Response rates to questionnaires were 73.7% (123/167) at 2 months, 65.9 (110/167) at 4 months, and 62.3% (104/167) at 6 months. Attrition was similar in all study groups. Among those with access to WebChoice, 64% (41/64) logged on more than once and 39% (25/64) sent e-messages to care providers. In the IPPC group, 40% (18/45) sent e-messages. Linear mixed models analyses revealed that the WebChoice group reported significantly lower symptom distress (mean difference 0.16, 95% CI 0.06-0.25, *P*=.001), anxiety (mean difference 0.79, 95% CI 0.09-1.49, *P*=.03), and depression (mean difference 0.79, 95% CI 0.09-1.49, *P*=.03) compared with the usual care group. The IPPC group reported significant lower depression scores compared with the usual care group (mean difference 0.69, 95% CI 0.05-1.32, *P*=.03), but no differences were observed for symptom distress or anxiety. No significant differences in self-efficacy were found among the study groups.

Conclusions: In spite of practice variations and moderate use of the interventions, our results suggest that offering Web support as part of regular care can be a powerful tool to help patients manage their illness. Our finding that a nurse-administered IPPC

XSL•FO
RenderX

alone can significantly reduce depression is particularly promising. However, the multicomponent intervention WebChoice had additional positive effects.

Trial Registration: Clinicaltrials.gov:NCT00971009; http://clinicaltrials.gov/show/NCT00971009 (Archived by WebCite at http://www.webcitation.org/6USKezP0Y).

(J Med Internet Res 2014;16(12):e295) doi:10.2196/jmir.3348

KEYWORDS

Web-based intervention; electronic mail; Internet; eHealth; cancer; patient-centered care; symptom management; professional-patient relations; randomized controlled trial

Introduction

The number of Web-based support systems to enhance self-management for people living with health conditions has increased rapidly in the last decade, and such interventions have been shown to assist a wide range of patients [1-11]. In cancer care, Web-based support systems are described as helpful for individuals [12]. This includes findings of increased health information competence [13,14], emotional processing [13], fighting spirit [15], social support [14,16], quality of life [16,17], as well as reductions in symptom distress [18,19], and decrease in depression and anxiety scores [17].

However, it can be difficult to distinguish which components of Web-based support systems are most beneficial for patients, and little is known about the relative contribution of different components [13]. In a study of a support system for cancer patients by Baker et al, different features of the system were tested and compared [13]. Results suggested that the benefit of the system was connected to the information (information about cancer, Web links, news, etc) and support services (support from peers and professionals), and that complex services such as coaching and tailoring of content did not produce benefits beyond simple access to the Internet. Another study of the same support system highlighted that benefits depend on *how* a patient uses a system, far more than the total amount of exposure or type of content that is chosen [20]. Overall time spent on the system showed no relation to outcomes. Improvement in patient status was connected to the commitment to use the system over time, independent of how much time they spent on the system. However, high use of communication services (discussion groups and “ask experts or peers”) were associated with decreased negative emotions. Knowledge of the use and effects of single components on patient outcomes will be important to determine component candidates for inclusion in Web-based support systems [13,16,21].

One component often offered as part of multicomponent systems, or as a standalone service, is e-messages. Several studies report benefits from using Internet-based patient-provider communication services (IPPC) for communication between patients and health care providers in terms of assisting patients in managing illness and improving health outcomes [22-24], addressing unmet communication needs [25,26], increasing satisfaction [23,27], and improving quality of care [22,27]. Most of these studies used IPPCs between patients and physicians. In an earlier study of WebChoice, the same Web-based support system for cancer patients used in this study, the nurse-administered IPPC was used by patients to ask questions and raise concerns related to symptom experiences, fear of relapses, and uncertainty in everyday life [18,28]. The IPPC was rated by patients as the most valuable component of WebChoice [29]. High levels of satisfaction with a nurse-administered IPPC were also reported in a study by Cornwall et al [30], but effects on patients’ outcomes of IPPCs alone are rarely described. Thus, we know little about the effect of standalone IPPCs and patient outcomes and how they compare to more comprehensive Web-based support systems where IPPCs are one of several components. In addition, as Web-based support systems require many resources in system development and updating compared to an IPPC, it is interesing to test the effects of these two intervention types compared to usual care.

WebChoice is a Web-based illness management support system based on patient-centered principles and designed to support cancer patients in self-management of their illness, independent of location and time [31]. The purpose of WebChoice is to help cancer patients reduce their symptom distress, improve emotional well-being, and enhance self-efficacy. Results from a previously randomized clinical trial (RCT) that followed 325 breast cancer and prostate cancer patients for 1 year showed that patients with access to WebChoice had significantly reduced symptom distress compared with the usual care control group [18]. Patients in the WebChoice group also had significant within-group improvements in depression during the study period. In addition, the control group experienced significant deterioration in self-efficacy and health-related quality of life during the study. One of the WebChoice features most valued by the study participants, as reported in the previous study, was the opportunity to send e-messages to expert nurses in cancer care, who responded to patients’ questions and concerns within 24 hours [29]. In the RCT described above, WebChoice was offered as a service to patients independent of location and clinical practice. The patients were recruited through advertisements and postal mail. The nurses who answered the e-messages had no treatment responsibilities and did not know the patients. Thus, so far we do not know if similar effects would be achieved if WebChoice or an IPPC were offered as an integrated part of regular care.

As several studies show benefits to Web-based support systems, it is timely to examine the relative contribution of different components of these multicomponent support systems aiming to improve selected outcomes. As cancer patients experience their illness primarily through symptoms, reduction of symptom distress, anxiety, and depression are important indicators of the success of illness management support. Furthermore, there is a

XSL•FO
RenderX

need to test the effects of Web-based support as a part of regular care.

The aims of the study were therefore to test and compare the effects of (1) an IPPC, (2) the multicomponent WebChoice intervention (including an IPPC), and (3) usual care (control group) on symptom distress, anxiety, and depression (primary outcomes), as well as self-efficacy (secondary outcome) after 6 months of follow-up. In addition, explorative sub-analyses were performed to detect whether the outcomes were associated with the actual use of the interventions. We hypothesized that the WebChoice group compared with the usual care group would have better outcomes than the IPPC group compared with the usual care group. We also hypothesized that both groups would have better outcomes than the usual care group, on both primary and secondary outcomes.

Methods

Subjects and Settings

We conducted an RCT with three groups: two intervention groups (the IPPC service and WebChoice) and a usual care group (clinical trial NCT00971009). Due to slower recruitment than anticipated, we had to stop inclusion after 200 consenting participants, before the calculated sample was obtained. The current paper reports on 167 patients for whom 6-month follow-up data were available at the time this paper was written. These participants were recruited between May 2010 and September 2012.

Inclusion criteria were recent diagnosis of breast cancer treated with surgery, or under treatment with radiation, chemotherapy, hormone therapy, or combinations of those (maximum 12 months after surgery), age over 18 years, able to write/read and speak Norwegian, having access to the Internet at home, and having a public key infrastructure (PKI) solution for secure system access.

Study participants were recruited from three hospitals in Norway—one university hospital and two regional hospitals—at breast diagnostic centers or the ambulatory chemotherapy, radiation, and surgical units. Participants did not receive any incentives for participating in the study.

See Multimedia Appendix 1 for the consent form and Multimedia Appendix 2 for the CONSORT-EHEALTH checklist [32].

Study Procedures

Eligible patients scheduled for surgery or coming in for checkups after surgery or treatment were identified by the study nurses at the hospitals and provided with information about the study. Upon patients' arrival at the clinic, the study nurses met the patients, provided brief information about the study, and asked if they were interested in participating. If the patients agreed, the nurse informed them about the study's purpose and procedures and asked for written informed consent. Consenting patients completed baseline questionnaires before randomization.

After completion of baseline questionnaires, patients were randomized according to a pre-defined automated computerized block randomization, with a block size of 42 stratified by site. Due to the content of the interventions, patients could not be blinded to which arm they were randomized.

Patients randomized to the usual care group were followed up as usual at the hospitals where they were treated. Patients who were randomized into the IPPC or WebChoice groups were informed and instructed in the use of the IPPC or WebChoice. They received a printed user manual with instructions for use, how to log on to the system, and an address and phone number to contact for help if needed. The only in-person information given was instructions on how to access the site and how to connect with the study support service if questions occurred. The study nurses showed them where in the user manual they could go to find information about how to access the site and how to connect with the study support service if needed. In addition, the participants were informed that they could use the IPPC or any component of WebChoice as much or as little as they liked and that using the system was entirely voluntary. The IPPC component was the same for both intervention groups.

After being informed about group assignment, the patients were given access to the interventions the same day. They received an automatic welcome message when the system was ready to use. There was an option to be notified by text message or regular email when new messages appeared in the system. Most participants wanted this notification. All participants were sent questionnaires by postal mail at 2, 4, and 6 months after enrolling.

In total, 20 care providers answered questions from patients: a dedicated group of expert nurses (n=11) and physicians (n=6) in breast cancer care, and social workers (n=3) at the hospital where the patients were treated. They were thoroughly trained in administering the IPPC, technically as well as in codes of conduct for online communication with patients. There was a clear schedule for who was responsible for answering patients' messages. The nurses were frontline and received all messages first. If necessary, they could forward the message to other care providers. If considered important, information from e-messages could be copied into the medical record and made available for other health care providers. When new questions arrived in the system, the recipient was notified through the hospital's email system or by text message. The same providers answered e-messages from both the IPPC and WebChoice groups using the same interface. However, they were not entirely blinded to the intervention group assignment because this was sometimes disclosed by patients through the messages. The health care providers had no access to details about how patients used other components of WebChoice. They did not receive any reimbursement or additional dedicated time for answering secure e-messages from the participants.

The study was approved by the Regional Committee for Medical and Health Research Ethics and the Data Security Inspectorate in Norway. Written informed consent was obtained from all participants. All data were submitted to a secure server using an encrypted connection. Patients and health care personnel were authenticated using a public key solution that is currently used by Norwegian banks as a security platform. This means the users' logon procedure is the same whether they log on to

XSL•FO
RenderX

their online bank or to IPPC/WebChoice. Thus, patients did not need to learn a new procedure.

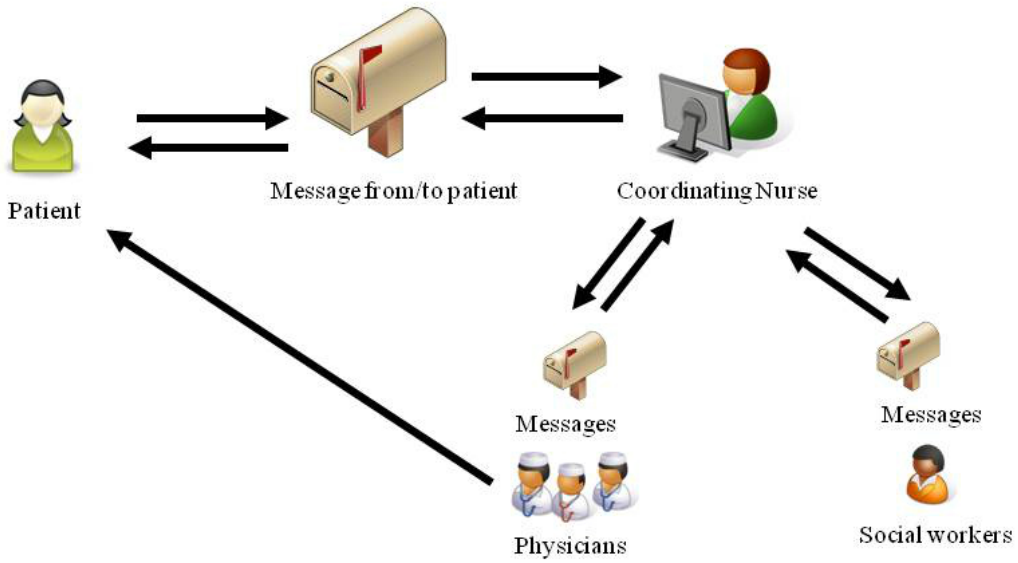
The system experienced a 9-day period of downtime during the first year due to technical problems at the hospital server that hosted the application. No changes were made to the interventions during the trial period apart from fixing minor bugs. The interventions could be used from different Internet browsers and were independent of Internet connection speed.

Description of the Two Interventions

Internet-Based Patient-Provider Communication Intervention

The IPPC is a further development of the IPPC component described in a previous study of WebChoice [18,33]. It allows patients to seek help from health care personnel at their treatment hospital. They can ask questions, share experiences with, or get advice from oncology nurses. If needed, the nurse can pass on their question to physicians and social workers (Figure 1). The system has a high security level, where both patients and health care providers log into the system with strong authentication keys. Care providers had access to the patients' medical records at the hospital. The patient questions were asynchronous and were answered within 2 work days (usually within 1 day).

Figure 1. Message flow between patients and health care providers.



WebChoice Intervention (Internet-Based Patient-Provider Communication Included)

WebChoice was developed in close cooperation with users and health care personnel [31]. After testing the system in an RCT [18], it was refined based on responses from users through questionnaires and focus groups. In addition, a blog feature was included. The WebChoice version tested in the current study targeted breast cancer patients and contained the following components in addition to the IPPC service [34] (Figure 2):

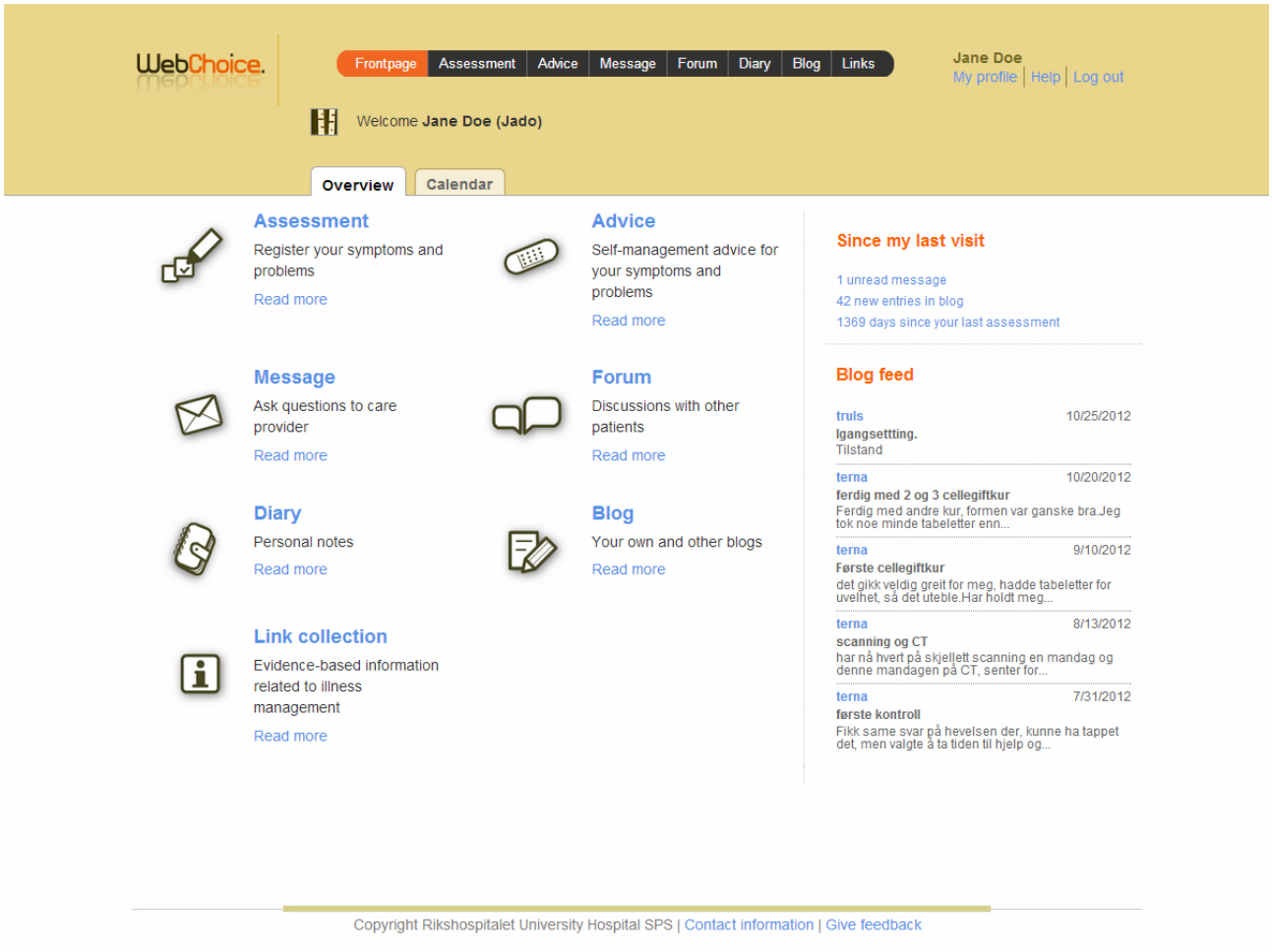
1. An assessment component where patients could monitor their symptoms, problems, and priorities for support along physical, functional, and psychosocial dimensions. From a predefined list, patients could choose symptoms and problems they were experiencing, rate the burden of these, and indicate where they needed help. This information could be used to monitor improvement/deterioration of the condition, indicate when to alert health care personnel, prepare for a hospital/physician consultation, improve patient-provider communication, or obtain immediate access to the self-management advice components described below.
2. An advice component provided illness self-management support. The patients' self-reported symptoms triggered the display of appropriate self-management activities that patients could choose from to relieve symptoms and problems. The component could also be used without finishing an assessment first. Each choice contained an explanation of what the activity was, how to perform it, potential risks, side effects, contraindications, when to contact a physician, levels of evidence, references to the source of the evidence, and links to other reliable websites for related information. The advice component was updated once a year.
3. An information component where patients had access to other reliable Web sources in Norwegian and English, such as information about tests, treatments and potential side effects, lifestyle suggestions, and information about patients' legal rights. External links were automatically checked every fourth week to ensure they were still active.
4. A communication component for sharing experiences with other patients. Patients could participate in an online forum group discussion that allowed them to exchange messages anonymously with other patients or use a blog. The forum and blog were monitored by nurses at our research center. The nurses did not participate in the forum or blog but

XSL•FO
RenderX

answered two postings in the forum not answered by the other participants.

5. An electronic diary where patients could keep personal notes.

Figure 2. Screenshot of the WebChoice overview page.



Measures and Data Collection

Overview

The primary outcomes were symptom distress, anxiety, and depression. The secondary outcome was self-efficacy. All outcomes were measured at baseline, 2, 4, and 6 months thorough self-assessed questionnaires sent to participants by postal mail.

Demographic Characteristics

Patients provided information on age, marital status, level of education, employment status, income, and the use of Internet services.

Patient Characteristics

The time of diagnosis and stage of disease were obtained from the medical record. Based on the patients' tumor (T), node (N), and metastasis (M) classification at the time of diagnosis, the stage of disease was classified into 5 stages (0=ductal carcinoma in situ, to 4=advanced-stage disease) using the TNM Classification of Malignant Tumours of the Union for International Cancer Control guidelines [35].

Comorbidity

Patients completed the Self-Administered Comorbidity Questionnaire (SCQ-19), which evaluated the number of, treatments for, and functional impact of health problems. It includes 16 common comorbidities and three optional conditions [36]. The total SCQ-19 score can range from 0-57 when the three optional items are used. It is a clinical scale, with established validity and reliability [36], for the assessment of comorbidities in patients with chronic medical conditions. A higher total score indicates a more severe comorbidity profile.

Symptom Distress

Symptom distress was measured by using the 32-item Memorial Symptom Assessment Scale (MSAS) [37], which lists physical and psychological symptoms that occur due to cancer or its treatment. For each symptom, patients were asked to indicate whether they had had the symptom during the previous week. If they had experienced the symptom, they were asked to rate its frequency, severity, and distress. Symptom frequency and severity was rated using a 4-point Likert scale. Symptom distress was rated using a 5-point Likert scale. The reliability and validity of the MSAS are well established [37], and MSAS has previously been used in breast cancer populations [38]. Higher

XSL•FO
RenderX

scores indicate greater symptom distress. Cronbach alpha coefficient for our sample at baseline was .85.

Anxiety and Depression

Anxiety and depression were measured with the Hospital Anxiety and Depression Scale (HADS) [39], a 14-item, self-report measure of psychological distress. The HADS is divided into 2 subscales: anxiety (HADS-A) (7 items) and depression (HADS-D) (7 items). Respondents are asked to indicate which of 4 response options (rated from 0-3; score range, 0-42) comes closest to describing how they have been feeling in the previous week for each item. Scores from 0-7 on the subscales are regarded as being in the normal range; a score of 11 or higher indicates a probable presence of a mood disorder, and a score of 8-10 is suggestive of the presence of the state [40]. The scale is found to perform well in assessing the symptom severity of anxiety disorders and depression in hospital settings, in primary health care and in the general population [41], and has demonstrated acceptable reliability in cancer populations [42]. Cronbach alpha coefficient is reported to vary from .68-.93 for HADS-A, and for HADS-D from .67-.90 [41]. In our sample, Cronbach alpha at baseline was .83 for HADS-A and .76 for HADS-D.

Self-Efficacy

Self-efficacy was measured with the Cancer Behavioral Inventory (CBI) version 2.0 [43], a 33-item instrument that measures coping self-efficacy with cancer-related stress on 7 dimensions: (1) maintenance of activity and independence, (2) seeking and understanding medical information, (3) stress management, (4) coping with treatment-related side effects, (5) accepting cancer and maintaining a positive attitude, (6) affective regulation, and (7) seeking support. Responses on 9-point Likert scales ranged from 1 (not at all confident) to 9 (totally confident). Higher scores indicated greater self-efficacy. CBI was used in a previous study testing WebChoice among breast and prostate cancer patients [18] and, according to Merluzzi et al, has good internal consistency with a Cronbach alpha coefficient reported of .94 [43]. Cronbach alpha coefficient for our sample at baseline was .96.

System Use

Data on system use were extracted from the user logs on the server. Information was collected on how many times the users had logged on and which components of WebChoice were accessed or used actively.

Analysis

Overview

Data on baseline characteristics are presented as medians and ranges for continuous variables and as proportions with percentages for categorical data. Differences between users and non-users were analyzed using the chi-square test for pairs of categorical variables. The Mann-Whitney-Wilcoxon test was used for continuous data with skewed distributions.

Effectiveness

For analysis of between-group differences in symptom distress, anxiety, and depression (primary outcomes) and self-efficacy (secondary outcome), linear mixed models (LMM) for repeated measures were fitted. A diagonal covariance structure was used to model dependencies among measurements on the same individual at different time points. Models for each outcome consisted of 3 effects: measurement occasion (time), interventions (WebChoice, IPPC, usual care), and the interaction of time and intervention. All measured time points of the outcome variables are considered and the LMM approach therefore adjusts for baseline differences. To test whether potential confounders impacted the results, LMM adjusted for variables such as site, age, marital status, education, time since diagnosis, stage of disease, and comorbidity were fitted. Compared to the unadjusted models, these adjusted models revealed even larger differences in favor of the intervention groups compared to the usual care group. Taking the limited sample size into account and aiming to avoid over fitting, only the results from the unadjusted models are presented. As no statistically significant differences were observed between the study groups on demographic and disease-related factors at baseline, these models were not further adjusted for the possible confounders. The authors are aware that this might underestimate the true differences between the groups. Analyses of primary and secondary outcomes were conducted on an intention-to-treat basis, including all participants in each group, independent of whether they were users or non-users of the interventions. The model parameters are estimated using the classical maximum likelihood approach. No imputation of missing data was necessary or performed, as the LMM uses all data available to estimate the covariance matrix and model the dependencies. The results are presented as *P* values for the overall effect of the variables when baseline score and all time points are included. Moreover, overall mean differences are presented, that is, the difference between groups adjusted for baseline scores and taking all time points into consideration. Report of overall mean differences was chosen as we were interested in differences between the groups over the entire 6-month period.

Explorative Sub-Analyses

In addition, explorative sub-analyses were performed to detect whether the outcomes were associated with the actual use of the interventions. LMM for repeated measures were fitted. Models for each outcome were fitted with three factors: measurement occasion (time), interventions (user/non-user of WebChoice and IPPC), and the interaction of time and intervention. Age was added as a covariate because age is known to be associated with use of Web-based tools [44,45].

Analyses were carried out using SPSS version 18.0. *P*-values <.05 were considered statistically significant, and all tests were two-sided.

Results

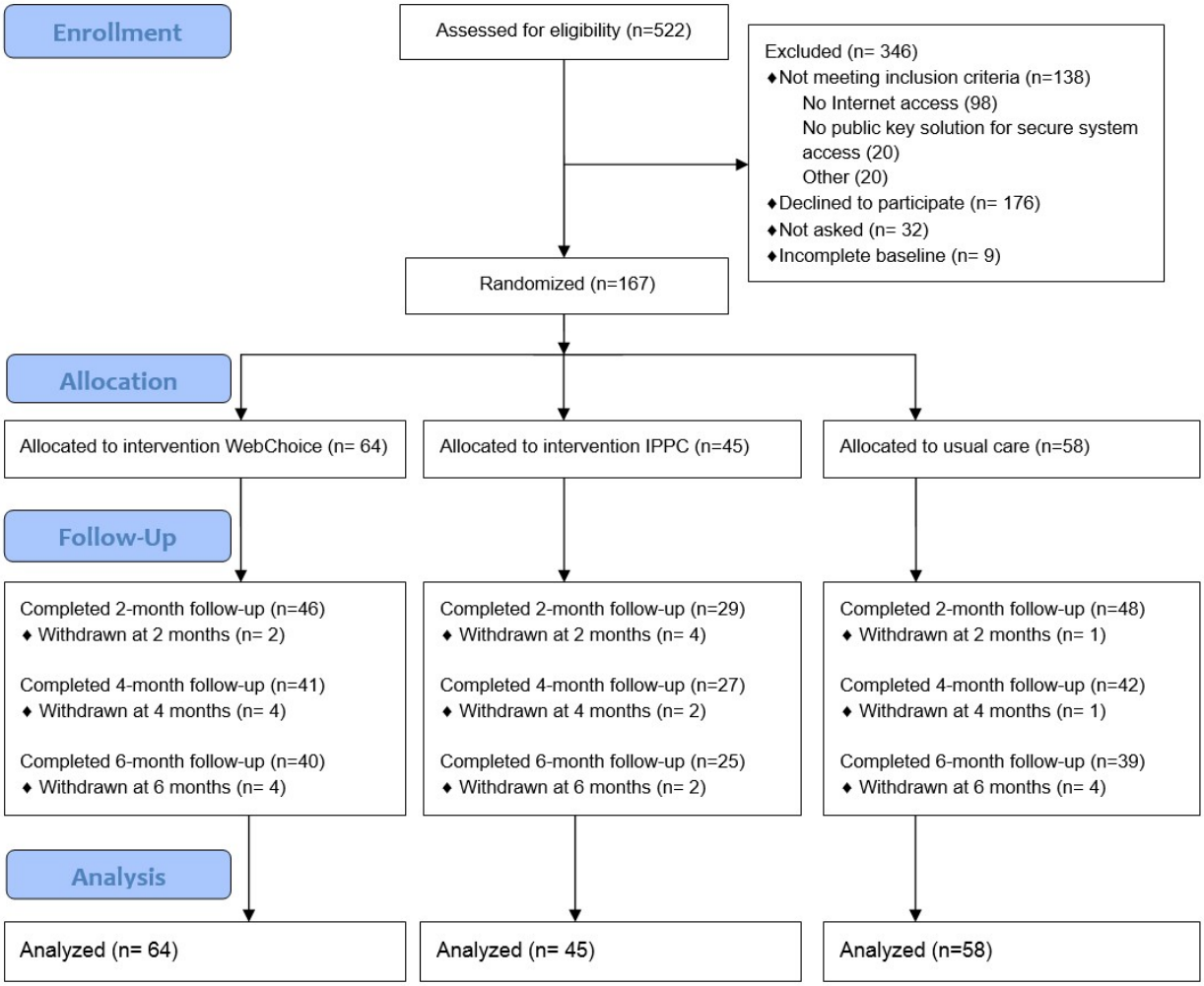
Participation Rates

The trial flow chart (Figure 3) shows the recruitment and retention at baseline and at 2-, 4- and 6-month follow-up. From May 2010 to September 2012, 522 breast cancer patients were assessed for eligibility. Of these, 138 did not meet the inclusion criteria, mostly due to lack of Internet access, and 176 declined

XSL•FO
RenderX

to participate. Non-participants were slightly older; the median age among those who did not meet the inclusion criteria was 67 years and among those who declined to participate, 59 years. Frequent reasons given for declining were lack of experience with computers/ Internet or that they had too much on their mind related to their illness. The 176 patients who agreed to participate in the study were randomized after filling in baseline questionnaires. Nine patients were excluded due to incomplete baseline data, leaving a sample of 167. During 6 months of follow-up, we had a 14% (23/167) attrition rate. Information on reasons for withdrawal is not available. There was no association between baseline characteristics and those who left the study. At the 6-month measurement, 62% (104/167) answered the questionnaires: WebChoice 62% (40/64), IPPC 57% (25/45), and usual care group 67% (39/58).

Figure 3. Flow of participants throughout the trial.



Baseline Characteristics

There were no statistically significant differences in demographics, disease-related factors, or outcome measures between participants in the three groups at baseline (Table 1).

XSL•FO
RenderX

Table 1. Baseline demographic and illness characteristics (n=167).

Characteristics	WebChoice (n=64)	IPPC (n=45)	Usual care (n=58)	*P* value
Age in years, median (range)	51 (37-79)	50 (31-66)	53 (36-69)	.21
Marital status, n (%)				.63
Married/cohabitating	52 (82)	34 (76)	43 (77)	
Single/divorced	11 (18)	11 (24)	13 (23)	
Education, n (%)				.29
Elementary/high school	23 (37)	22 (49)	30 (54)	
University/college ≤4 years	24 (39)	14 (31)	12 (21)	
University/college >4 years	15 (24)	9 (20)	14 (25)	
Household annual income, (NOK [a]), n (%)				.46
<400,000	8 (13)	12 (27)	10 (18)	
400,000 to 600,000	14 (23)	9 (20)	18 (32)	
600,000 to 800,000	19 (31)	11 (24)	13 (23)	
>800,000	21 (34)	13 (29)	15 (27)	
Employment status, n (%)				.54
Full-time/part-time work	21 (33)	18 (41)	14 (26)	
Sick leave/disability benefits	34 (54)	22 (50)	32 (58)	
Retired/other	8 (13)	4 (9)	9 (16)	
Stage of disease, n (%)				.53
0	2 (3)	1 (2)	5 (9)	
1	24 (38)	21 (47)	26 (45)	
2	32 (50)	20 (47)	24 (41)	
3	6 (9)	2 (4)	3 (5)	
Psychosocial factors, median (range)				
Symptom distress	.47 (.03-1.49)	.52 (.03-1.76)	.59 (.05-1.86)	.36
Anxiety	4.0 (0-16)	5.0 (0-17)	4.0 (0-15)	.38
Depression	1.5 (0-11)	2.0 (0-10)	2.0 (0-14)	.37
Self-efficacy	247 (110-297)	214 (101-297)	238 (102-297)	.11
Months since diagnosis, median (range)	1 (0-10)	0 (0-9)	1 (0-10)	.16
Comorbidity, median (range)	2 (0-10)	2 (0-16)	2.0 (0-13)	.51
Sending/receiving email, n (%)				
> 1x/week	57 (91)	37 (82)	53 (95)	.09
< 1x/week	2 (3)	6 (13)	3 (5)	
< 1x/month	4 (6)	2 (4)	0 (0)	
Reading information on Internet, n (%)				.28
> 1x/week	54 (87)	36 (80)	43 (78)	
< 1x/week	3 (5)	2 (4)	7 (13)	
< 1x/month	5 (8)	7 (16)	5 (9)	
Participation in social media/groups, n (%)				.67
> 1x/week	28 (45)	18 (40)	20 (36)	
< 1x/week	7 (11)	9 (20)	8 (15)	
Never	27 (44)	18 (40)	27 (49)	

XSL•FO
RenderX

[a]NOK=Norwegian kroner: 400,000≈$US 67,000; 600,000≈$US 100,000.

Effectiveness

Figure 4 and Table 2 provide the patients' self-reported scores for the primary outcomes of symptom distress, anxiety, and depression. When measurements at all four time points were included in the model, the WebChoice group reported significantly lower symptom distress over time (mean difference -0.16, 95% CI -0.25 to -0.06, *P*=.001), anxiety (mean difference -0.79, 95% CI -1.49 to -0.09, *P*=.03), and depression (mean difference -0.79, 95% CI 1.18 to -0.05, *P*=.03) compared with the usual care group. Over time, the IPPC group had significantly lower depression scores compared with the usual care group (mean difference -0.69, 95% CI -1.32 to -0.05, *P*=.03), but no differences were observed for symptom distress or anxiety. Time by condition interactions was examined, but no significant results detected. There were no statistically significant differences over time between the two intervention groups on symptom distress, anxiety, and depression (data not shown).

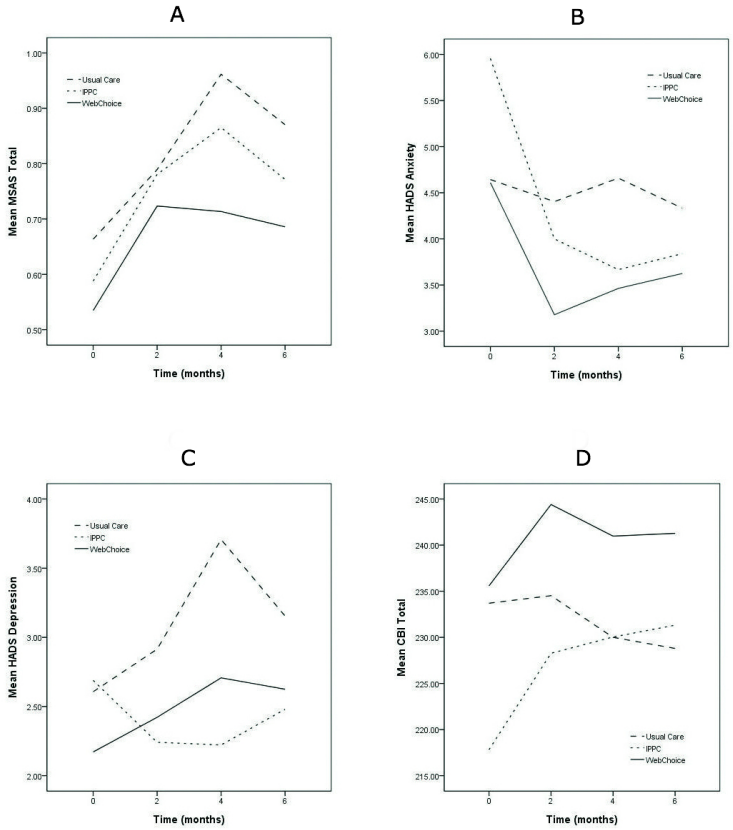
The WebChoice group tended to score higher than the usual care group on self-efficacy (secondary outcome) over time (mean difference 8.81, 95% CI -0.92 to 18.53, *P*=.08) (Table 2 and Figure 4). No statistically significant differences were found over time between the IPPC group and the usual care group on self-efficacy.

Table 2. Effect of IPPC (n=45) and WebChoice (n=64) compared with the usual care group (n=58) on outcome variables: summary of basic mixed models.

	WebChoice			IPPC		
	Mean diff[a]	95% CI	*P* value	Mean diff[a]	95% CI	*P* value
MSAS total	-0.16	-0.25 to -0.06	.001	-0.07	-0.18 to 0.04	.21
HADS-A	-0.79	-1.49 to -0.09	.03	-0.14	-0.93 to 0.64	.72
HADS-D	-0.61	-1.18 to -0.05	.03	-0.69	-1.32 to -0.05	.03
CBI	8.81	-0.92 to 18.53	.08	-4.89	-15.90 to 6.12	.38

[a]Based on estimated marginal means. Analyses adjusted for baseline scores.

Figure 4. Estimated marginal means of (A) symptom distress (MSAS Total), (B) anxiety (HADS Anxiety), (C) depression (HADS Depression), and (D) self-efficacy (CBI Total) for the usual care group (n=58), the IPPC group (n=45), and the WebChoice group (n=64).



XSL•FO
RenderX

System Use

In accordance with the CONSORT-EHEALTH checklist [32], we analyzed the use of IPPC and WebChoice among those who had access to the applications. Among those randomized to the IPPC, 40% (18/45) sent at least an e-message and were defined as users (Table 3). In the WebChoice group, 77% (49/64) logged on at least once during the 6 months, and 39% (25/64) sent e-messages. We defined the two-thirds (41/64) who accessed WebChoice more than once as users. We required patients to log in at least twice before we defined them as users because patients who logged in only once may have read only the welcome message and never actually used the system. The users logged on a median 7 times (range 2-41).

Table 3 shows the use of the IPPC and the usage of different components in WebChoice among the users. Of the WebChoice users, 61% (25/41) sent e-messages, 20% (8/41) posted in the discussion forum, and 37% (15/41) posted their own blogs. However, patients visited the IPPC, forum, and blog more often to read information without submitting their own postings. Reading of other blogs was the component most highly accessed (total 684; range 0-58, median 14). The advice components were also highly accessed (total 317; range 0-62; median 5).

The IPPC messages were mainly answered by nurses within their regular working hours. Of the 153 messages sent by the respondents, 22% (33/153) were passed on to and answered by physicians. Only one message was passed on to social workers. Patient questions posed through the IPPC were always answered via secure e-messages. In a few instances, the nurses informed the patient that they could contact her by phone to discuss issues voiced in the message. Time spent on answering messages was not measured. However, there were no indications that the nurses or physician felt the e-message answering task had been too time consuming for them during the study.

There were no differences between users and non-users in either intervention group on any demographic or disease-related variables.

Table 3. Usage of components in WebChoice (n=41) and IPPC (n=18) over 6 months of access.

Components	Users WebChoice (n=41)					Users IPPC (n=18)				
	Times accessed			Users who accessed at least once		Times accessed			Users who accessed at least once	
	Median	IQR[a]	Range	n	%	Median	IQR[a]	Range	n	%
Total visits	7	11	2-41	41	100	9.5	10	1-17	18	100
E-messages sent	1	3	0-9	25	61	3.5	4	1-7	18	100
Total e-messages visits	4	7	0-25	39	95	12	12	2-24	18	100
Assessments	1	2	0-7	30	77					
Assessment visits	5	6	0-25	40	98					
Advice views	5	10	0-62	29	71					
Information section visits	2	3	0-28	34	83					
Posts in forum	0	0	0-3	8	20					
Forum visits	4	4	0-25	38	83					
Posts in blog	0	2	0-9	15	37					
View of others' blogs	14	19	0-58	37	90					
Diary notes	0	3	0-14	19	46					
Diary visits	3	6	0-26	35	85					

[a]IQR=interquartile range.

Exploratory Analysis: Effect of System Use on Outcomes

No differences were detected between users and non-users of WebChoice or IPPC on symptom distress, depression, or self-efficacy (data not shown). The users of the IPPC had significantly lower scores on anxiety compared with the non-users (mean difference -1.28, 95% CI -2.54 to -0.01, *P*=.047). No such differences between users and non-users were observed in the WebChoice group (data not shown).

Discussion

Principal Results

The current effectiveness study demonstrates that access to the multicomponent Web-based support system WebChoice for 6 months, among women with breast cancer within the first year after diagnosis, reduced symptom distress and levels of anxiety and depression scores. A tendency towards increased self-efficacy could also be detected for the WebChoice group. This is promising given three diverse practice settings in regular care. Also noteworthy is the finding that access to an IPPC alone reduced depression scores. These results support the hypothesis of the WebChoice group having better outcomes than the IPPC

XSL•FO
RenderX

group in symptom distress and anxiety, compared to the usual care group. The IPPC group had a similar effect on reduction of depression as WebChoice, however, and the WebChoice group did not have better self-efficacy than the IPPC group compared to usual care.

Our findings are in line with previous research showing that Web-based support systems in cancer populations can decrease depression and anxiety scores [17] and reduce symptom distress [18,19]. The reported effects in our study were detected despite a smaller sample size than initially calculated. However, in this study participants were more homogenous compared with the previously published study [18] that included both women with breast cancer and men with prostate cancer. In addition, the previous study included patients with recurrence of disease and metastasis, and patients were included independent of time since diagnosis (mean time since diagnosis 2.2 years), whereas patients in our study were included within the first year after diagnosis (median time since diagnosis 1 month).

One possible explanation for our significant results might be that both WebChoice and IPPC might meet unmet needs reported by cancer patients, such as needs within the communication, information, psychosocial, psychological, and supportive care domains, generally highest during the treatment phase [46]. WebChoice allows patients to monitor their psychological, psychosocial, and physical symptoms and also get individually tailored information and support on how to manage their symptoms through the advice component. The information component can offer educational information through access to other reliable Web sources. Through WebChoice, patients are also able to read the information repeatedly when it suits them. Other studies suggests that Web-based information sources are used for different purposes [47] and that it can be easier to address sensitive information through email service than in personal encounters with health care professionals [25]. Offering self-management interventions in this early phase might be especially helpful. The need for early interventions was also supported by a study of an IPPC service similar to the one in the current study, where the need for such a service was described as being most prominent during the first phase after discharge from hospital [25].

Our results indicating that the IPPC reduced depression scores are especially promising, as depression is one of the most debilitating symptoms people can have and is highly prevalent among cancer patients [48,49]. This result was detected despite respondents being recruited from different settings with variations in organization of care, which holds promise for IPPCs as effective interventions for reduction of depression scores among breast cancer patients across settings.

The IPPC's ability to reduce depression scores might relate to indications that patients with higher scores of depression, in addition to higher symptom distress and low social support, are high users of the IPPC service in WebChoice [47]. As such, the IPPC is an intervention that is used by those with high illness burden, a group with high needs, and potential for improvement. The median baseline scores of both anxiety and depression were below the defined cut-off scores of 8, which is predictive for presence of anxiety and depression [40,41]. However, reduction of scores indicates that people are feeling better. Our limited sample size did not allow analyses on how the intervention affected the individuals who had scores of 8 or more.

Whether the differences in effects of the IPPC and WebChoice are related to the additional features contained in WebChoice is not clear. The IPPC feature might be used differently in the two groups, and as such we were not able to fully disentangle the effect of the IPPC feature. The patients would have had to be allocated to separate conditions in order to test different features. However, as in the IPPC group where 40% (18/45) sent e-messages, 39 % (25/64) in the WebChoice group sent e-messages, suggesting that interest in using this component is the same, independent of other features available.

Use and Relation to Effects

Use was not connected to the observed effects on symptom distress and depression in the current study. Similar results were observed in the previous study of WebChoice [18]. The findings that IPPC users had significantly lower scores on anxiety (P=.047) than the non-users must be interpreted with caution. The groups compared were small (18 users and 27 non-users), and thus the finding might be spurious. WebChoice is an illness management support system designed to support cancer patients in self-management of their illness. The system offers different components, so participants can use what they prefer without any push to use the system. Our findings of positive effects in the study despite user frequencies of just 64% (WebChoice) and 40% (IPPC) might relate to the psychological effects of the sheer possibility of using the system when needed. Interviews with non-users of a similar IPPC revealed that even if they did not use the system, they liked having the possibility [50]. The assurance that someone is available and can answer the questions important to a patient may contribute to the effects observed on depression in the current study. The opportunity to get the information needed for self-management of symptoms and problems, independent of time and location, might also contribute to the other findings of reduced symptom distress and anxiety.

Limitations

Several limitations need to be addressed. The first limitation concerns the small sample size. A larger sample would have increased the validity of the study but would also have prolonged the time needed for recruitment. In the recruitment process, most of the potential participants were approached. However, only a third of those approached were included. Lack of access to Internet, the most frequent reason for not meeting the inclusion criteria, was reported by 19% (98/522) of those approached. Among those who were eligible, a frequent reason given for declining participation was that patients judged their computer and Internet skills as poor. One way to increase the participation rates might be to offer a demonstration of the interventions at the time of inclusion.

A smaller sample size than initially calculated and the attrition rate during the study reduced statistical power for our analyses. Because we had to stop inclusion of participants before the a priori calculated sample was obtained, block randomization led to different sample sizes in the three groups, with the least

XSL•FO
RenderX

participants in the IPPC group. The project was also subject to high attrition during the study, which is not uncommon in studies of eHealth interventions [51]. In addition, the IPPC group had the lowest number of completers of questionnaires at 6 months. Our analysis should thus be interpreted with caution, and additional research is needed to confirm our results.

Another limitation relates to low use of the interventions. The analyses of the intervention groups compared to the usual care group therefore compare the effects of a little used intervention. However, as the effects were detected through intention to treat analysis, the effect might be connected to the option and possibility of using the system, not necessarily to actual use. Our post hoc analyses of usage and its relation to outcomes were based on a smaller sample, comparing users and non-users, and must be viewed as an exploratory analysis only.

Interventions such as WebChoice, which offer components where people interact with each other (forum and blog), need a critical mass in order to be fully utilized. A study period of 6 months, an inclusion period of nearly 2.5 years, and the three-armed study design resulted in few participants receiving the WebChoice intervention simultaneously. This could be an explanation for low use of the discussion forum in this study.

Patients included in the study were younger (median age 52) than those who were excluded (median age 67) or who declined to participate (median age 59). In addition, the participants had higher education levels than the average level of education in Norway, suggesting that they were not representative of all age and educational groups among breast cancer patients. This, together with the small sample size, rate of declining participation, and the attrition from the study likely lowers the generalizability of the findings. On the other hand, the patients were recruited from three different hospitals across the country, which increased generalizability of findings across practice settings. Finally, nearly 50% of the participants were included right after the time of diagnosis. However, the other half was included at 1-10 months after diagnosis. As such, they were in different phases in their treatments and experienced different side effects at the time of questionnaire completion. This might have influenced the symptoms reported on the MSAS and HADS. On the other hand, this variability might strengthen our results.

Implications

This study illustrates the feasibility of offering parts of Web-based support systems in regular care, as the IPPC components were answered by nurses/physicians at the hospital where the patients were treated, providing the patients with easy access to the expertise, without a face-to-face appointment. The IPPC service, with its ability to reduce depression scores, will be an important component to integrate in Web-based support systems and can also be offered as a standalone system.

The integration of Web-based support into clinical practice will require some changes, and changing routines in care is challenging [52-54]. There are reports of skepticism among care providers about use of IPPCs in routine care [55,56]. Some health care providers have expressed concerns that the use of e-messages might disrupt existing workflows and increase workloads [55]. Patients, however, expect to be able to communicate with their health care providers through e-messages [27,55,57], and integration of the IPPC (as a standalone service or as part of multicomponent support systems) does not require a huge change in health care routines. In this study, only a few nurses and physicians were trained to answer the IPPC. The number of e-messages in the study was moderate and most were answered by nurses, during their regular working hours. As such, the IPPC did not interfere with the workflow of the entire staff, and the number of e-messages was reported as manageable. The nurses answering the IPPC performed their new task during regular working hours, without any incentives. This indicates that it is feasible to implement IPPCs in regular care and that the service can be managed and answered by nurses. If clinicians recognize Web-based support as effective and easy-to-access resources for their patients' self-management support and outcomes, they might be more receptive to these types of interventions or added service options [58].

In the current study, the e-messages were primarily answered by nurses and passed on to physicians only if needed, which indicates that the IPPC can successfully be managed by nurses on the front line. Nurses are described as having a holistic approach to patients, focusing on emotional issues, consequences of disease, and illness information [59]. They are thus well equipped to answer questions and concerns and are reported to be sensitive to and able to respond to patients' emotions expressed through e-messages [60].

WebChoice, with all its features, had added effects compared with the IPPC alone. However, development and updating of systems such as WebChoice require far more resources than an IPPC service alone. Furthermore, IPPCs can be used for different patient groups independent of diagnosis, as the patient and provider are the ones who create the content.

Further Research

An aspect that remains to be tested is whether Web-based support systems are more effective when health care personnel with treatment responsibilities for the patients answer messages within the system, rather than health care personnel without this knowledge (patients can send e-messages anonymously). To obtain a deeper understanding of experience with the use of IPPC in routine care, we are currently interviewing nurses and physicians who have answered e-messages in this study. This experience is important to guide the implementation processes in the future. In addition, as most studies report on services between patients and physicians, more research is needed to test similar services managed by nurses. Finally, the positive effects on patients' outcomes, despite moderate user frequencies and almost no differences detected between users and non-users, calls for further research examining how the psychological effect of simply having access to information and support might impact outcomes.

Conclusions

In spite of the practice variations at three different hospitals, and moderate use of the IPPC service and WebChoice by study patients, our preliminary results suggest that offering Web-based

XSL•FO
RenderX

support as a part of regular practice can be a powerful tool to help patients manage their illness. Our finding that the nurse-administered IPPC alone significantly reduced depression, a highly debilitating symptom, is particularly promising, as an IPPC can be implemented in different settings. However, the multicomponent support system, WebChoice, had additional positive effects on reducing anxiety symptoms and symptom distress.

Acknowledgments

The study was funded by the South-East Regional Health Authority of Norway (grant number 2009051). WebChoice was provided by the Centre for Shared Decision Making and Collaborative Care Research at Oslo University Hospital, Norway. The center delivered the intervention and managed the trial.

We would like to thank all patients, nurses, physicians, and social care workers who participated in the study.

Conflicts of Interest

The last author (CR) is the developer of WebChoice but has no financial or ownership rights to the application. EB, MC, ME, and CR are affiliated with the developing center but have no financial or non-financial interest to declare in relation to the study.

Multimedia Appendix 1

Informed consent.

[PDF File (Adobe PDF File), 109KB - jmir_v16i12e295_app1.pdf]

Multimedia Appendix 2

CONSORT-EHEALTH checklist V1.6.2 [32].

[PDF File (Adobe PDF File), 1001KB - jmir_v16i12e295_app2.pdf]

References

1. Kuijpers W, Groen WG, Aaronson NK, van Harten WH. A systematic review of web-based interventions for patient empowerment and physical activity in chronic diseases: relevance for cancer survivors. J Med Internet Res 2013;15(2):e37 [FREE Full text] [doi: 10.2196/jmir.2281] [Medline: 23425685]
2. Solomon M, Wagner SL, Goes J. Effects of a Web-based intervention for adults with chronic conditions on patient activation: online randomized controlled trial. J Med Internet Res 2012;14(1):e32 [FREE Full text] [doi: 10.2196/jmir.1924] [Medline: 22353433]
3. Solomon MR. Information Technology to Support Self-Management in Chronic Care. Disease Management & Health Outcomes 2008;16(6):391-401. [doi: 10.2165/0115677-200816060-00004]
4. Murray E, Burns J, See TS, Lai R, Nazareth I. Interactive Health Communication Applications for people with chronic disease. Cochrane Database Syst Rev 2005(4):CD004274. [doi: 10.1002/14651858.CD004274.pub4] [Medline: 16235356]
5. Glasgow RE, Kurz D, King D, Dickman JM, Faber AJ, Halterman E, et al. Twelve-month outcomes of an Internet-based diabetes self-management support program. Patient Educ Couns 2012 Apr;87(1):81-92 [FREE Full text] [doi: 10.1016/j.pec.2011.07.024] [Medline: 21924576]
6. Glasgow RE, Kurz D, King D, Dickman JM, Faber AJ, Halterman E, et al. Outcomes of minimal and moderate support versions of an internet-based diabetes self-management support program. J Gen Intern Med 2010 Dec;25(12):1315-1322 [FREE Full text] [doi: 10.1007/s11606-010-1480-0] [Medline: 20714820]
7. Lorig K, Ritter PL, Laurent DD, Plant K, Green M, Jernigan VB, et al. Online diabetes self-management program: a randomized study. Diabetes Care 2010 Jun;33(6):1275-1281 [FREE Full text] [doi: 10.2337/dc09-2153] [Medline: 20299481]
8. Lorig KR, Ritter PL, Laurent DD, Plant K. The internet-based arthritis self-management program: a one-year randomized trial for patients with arthritis or fibromyalgia. Arthritis Rheum 2008 Jul 15;59(7):1009-1017 [FREE Full text] [doi: 10.1002/art.23817] [Medline: 18576310]
9. Powell J, Hamborg T, Stallard N, Burls A, McSorley J, Bennett K, et al. Effectiveness of a web-based cognitive-behavioral tool to improve mental well-being in the general population: randomized controlled trial. J Med Internet Res 2013;15(1):e2 [FREE Full text] [doi: 10.2196/jmir.2240] [Medline: 23302475]
10. van Straten A, Cuijpers P, Smits N. Effectiveness of a web-based self-help intervention for symptoms of depression, anxiety, and stress: randomized controlled trial. J Med Internet Res 2008;10(1):e7 [FREE Full text] [doi: 10.2196/jmir.954] [Medline: 18364344]
11. Stellefson M, Chaney B, Barry AE, Chavarria E, Tennant B, Walsh-Childers K, et al. Web 2.0 chronic disease self-management for older adults: a systematic review. J Med Internet Res 2013;15(2):e35 [FREE Full text] [doi: 10.2196/jmir.2439] [Medline: 23410671]

XSL•FO
RenderX

12. Ventura F, Ohlén J, Koinberg I. An integrative review of supportive e-health programs in cancer care. Eur J Oncol Nurs 2013 Aug;17(4):498-507. [doi: 10.1016/j.ejon.2012.10.007] [Medline: 23158437]
13. Baker TB, Hawkins R, Pingree S, Roberts LJ, McDowell HE, Shaw BR, et al. Optimizing eHealth breast cancer interventions: which types of eHealth services are effective? Transl Behav Med 2011 Mar;1(1):134-145 [FREE Full text] [Medline: 21709810]
14. Gustafson DH, McTavish FM, Stengle W, Ballard D, Hawkins R, Shaw BR, et al. Use and Impact of eHealth System by Low-income Women With Breast Cancer. J Health Commun 2005;10 Suppl 1:195-218. [doi: 10.1080/10810730500263257] [Medline: 16377608]
15. David N, Schlenker P, Prudlo U, Larbig W. Internet-based program for coping with cancer: a randomized controlled trial with hematologic cancer patients. Psychooncology 2013 May;22(5):1064-1072. [doi: 10.1002/pon.3104] [Medline: 22565413]
16. Gustafson DH, Hawkins R, McTavish F, Pingree S, Chen WC, Volrathongchai K, et al. Internet-Based Interactive Support for Cancer Patients: Are Integrated Systems Better? J Commun 2008 Jun;58(2):238-257 [FREE Full text] [doi: 10.1111/j.1460-2466.2008.00383.x] [Medline: 21804645]
17. Yun YH, Lee KS, Kim YW, Park SY, Lee ES, Noh DY, et al. Web-based tailored education program for disease-free cancer survivors with cancer-related fatigue: a randomized controlled trial. J Clin Oncol 2012 Apr 20;30(12):1296-1303 [FREE Full text] [doi: 10.1200/JCO.2011.37.2979] [Medline: 22412149]
18. Ruland CM, Andersen T, Jeneson A, Moore S, Grimsbø G, Børøsund E, et al. Effects of an internet support system to assist cancer patients in reducing symptom distress: a randomized controlled trial. Cancer Nurs 2013;36(1):6-17. [doi: 10.1097/NCC.0b013e31824d90d4] [Medline: 22495503]
19. Gustafson DH, DuBenske LL, Namkoong K, Hawkins R, Chih MY, Atwood AK, et al. An eHealth system supporting palliative care for patients with non-small cell lung cancer: a randomized trial. Cancer 2013 May 1;119(9):1744-1751 [FREE Full text] [doi: 10.1002/cncr.27939] [Medline: 23355273]
20. Han JY, Hawkins RP, Shaw BR, Pingree S, McTavish F, Gustafson DH. Unraveling Uses and Effects of an Interactive Health Communication System. J Broadcast Electron Media 2009;53(1):112-133 [FREE Full text] [doi: 10.1080/08838150802643787] [Medline: 23172985]
21. Hawkins RP, Pingree S, Baker T, Roberts LJ, Shaw B, McDowell H, et al. Integrating eHealth With Human Services for Breast Cancer Patients. Transl Behav Med 2011 Mar;1(1):146-154 [FREE Full text] [doi: 10.1007/s13142-011-0027-1] [Medline: 22496712]
22. Zhou YY, Kanter MH, Wang JJ, Garrido T. Improved quality at Kaiser Permanente through e-mail between physicians and patients. Health Aff (Millwood) 2010 Jul;29(7):1370-1375 [FREE Full text] [doi: 10.1377/hlthaff.2010.0048] [Medline: 20606190]
23. Goldzweig LC, Shekelle PG, Towfigh AA, Paige NM, Orshansky G, Haggstrom DA. Systematic review: Secure messaging between providers and patients, and patients' access to their own medical record. Evidence-based Synthesis Program 2012 Jul;39(1) [FREE Full text] [Medline: 22973584]
24. de Jong CC, Ros WJ, Schrijvers G. The effects on health behavior and health outcomes of Internet-based asynchronous communication between health providers and patients with a chronic condition: a systematic review. J Med Internet Res 2014;16(1):e19 [FREE Full text] [doi: 10.2196/jmir.3000] [Medline: 24434570]
25. Wibe T, Hellesø R, Varsi C, Ruland C, Ekstedt M. How does an online patient-nurse communication service meet the information needs of men with recently diagnosed testicular cancer? ISRN Nurs 2012;2012:260975 [FREE Full text] [doi: 10.5402/2012/260975] [Medline: 23251816]
26. Ye J, Rust G, Fry-Johnson Y, Strothers H. E-mail in patient-provider communication: a systematic review. Patient Educ Couns 2010 Aug;80(2):266-273 [FREE Full text] [doi: 10.1016/j.pec.2009.09.038] [Medline: 19914022]
27. McGeady D, Kujala J, Ilvonen K. The impact of patient-physician web messaging on healthcare service provision. Int J Med Inform 2008 Jan;77(1):17-23. [doi: 10.1016/j.ijmedinf.2006.11.004] [Medline: 17188564]
28. Grimsbø G, Finset A, Ruland CM. Left hanging in the air: experiences of living with cancer as expressed through E-mail communications with oncology nurses. Cancer Nurs 2011;34(2):107-116. [doi: 10.1097/NCC.0b013e3181eff008] [Medline: 20921887]
29. Ruland CM, Maffei RM, Børøsund E, Krahn A, Andersen T, Grimsbø G. Evaluation of different features of an eHealth application for personalized illness management support: cancer patients' use and appraisal of usefulness. Int J Med Inform 2013 Jul;82(7):593-603. [doi: 10.1016/j.ijmedinf.2013.02.007] [Medline: 23507561]
30. Cornwall A, Moore S, Plant H. Embracing technology: patients', family members' and nurse specialists' experience of communicating using e-mail. Eur J Oncol Nurs 2008 Jul;12(3):198-208. [doi: 10.1016/j.ejon.2007.09.008] [Medline: 18406667]
31. Ruland CM, Jeneson A, Andersen T, Andersen R, Slaughter L, Bente-Schjødt-Osmo, et al. Designing tailored Internet support to assist cancer patients in illness management. AMIA Annu Symp Proc 2007:635-639 [FREE Full text] [Medline: 18693913]
32. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

XSL•FO
RenderX

33. Ruland CM, Borosund E, Varsi C. User requirements for a practice-integrated nurse-administered online communication service for cancer patients. Stud Health Technol Inform 2009;146:221-225. [Medline: 19592838]
34. Oslo University Hospital. WebChoice 2. Centre for Shared Decision Making an Collaborative Care Research URL: http://www.communicaretools.org/tools/the-webchoice-tool/ [accessed 2014-08-22] [WebCite Cache ID 6S0rdp7AA]
35. International Union Against Cancer. In: Sobin LH, Gospodarowicz M, Wittekind C, editors. TNM Classification of Malignant Tumours. Oxford, UK: Wiley-Blackwell; 2010.
36. Sangha O, Stucki G, Liang MH, Fossel AH, Katz JN. The Self-Administered Comorbidity Questionnaire: a new method to assess comorbidity for clinical and health services research. Arthritis Rheum 2003 Apr 15;49(2):156-163 [FREE Full text] [doi: 10.1002/art.10993] [Medline: 12687505]
37. Portenoy RK, Thaler HT, Kornblith AB, Lepore JM, Friedlander-Klar H, Kiyasu E, et al. The Memorial Symptom Assessment Scale: an instrument for the evaluation of symptom prevalence, characteristics and distress. Eur J Cancer 1994;30A(9):1326-1336. [Medline: 7999421]
38. Hofsø K, Rustøen T, Cooper BA, Bjordal K, Miaskowski C. Changes over time in occurrence, severity, and distress of common symptoms during and after radiation therapy for breast cancer. J Pain Symptom Manage 2013 Jun;45(6):980-1006. [doi: 10.1016/j.jpainsymman.2012.06.003] [Medline: 23026547]
39. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983 Jun;67(6):361-370. [Medline: 6880820]
40. Snaith RP. The Hospital Anxiety And Depression Scale. Health Qual Life Outcomes 2003;1(29):- [FREE Full text] [doi: 10.1186/1477-7525-1-29] [Medline: 12914662]
41. Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale. An updated literature review. J Psychosom Res 2002 Feb;52(2):69-77. [Medline: 11832252]
42. Alexander S, Palmer C, Stone PC. Evaluation of screening instruments for depression and anxiety in breast cancer survivors. Breast Cancer Res Treat 2010 Jul;122(2):573-578. [doi: 10.1007/s10549-009-0669-6] [Medline: 19960243]
43. Merluzzi TV, Nairn RC, Hegde K, Martinez Sanchez MA, Dunn L. Self-efficacy for coping with cancer: revision of the Cancer Behavior Inventory (version 2.0). Psychooncology 2001;10(3):206-217 [FREE Full text] [Medline: 11351373]
44. Kelders SM, Van Gemert-Pijnen JE, Werkman A, Nijland N, Seydel ER. Effectiveness of a Web-based intervention aimed at healthy dietary and physical activity behavior: a randomized controlled trial about users and usage. J Med Internet Res 2011;13(2):e32 [FREE Full text] [doi: 10.2196/jmir.1624] [Medline: 21493191]
45. Kerr C, Murray E, Noble L, Morris R, Bottomley C, Stevenson F, et al. The potential of Web-based interventions for heart disease self-management: a mixed methods investigation. J Med Internet Res 2010;12(4):e56 [FREE Full text] [doi: 10.2196/jmir.1438] [Medline: 21156471]
46. Harrison JD, Young JM, Price MA, Butow PN, Solomon MJ. What are the unmet supportive care needs of people with cancer? A systematic review. Support Care Cancer 2009 Aug;17(8):1117-1128. [doi: 10.1007/s00520-009-0615-5] [Medline: 19319577]
47. Børøsund E, Cvancarova M, Ekstedt M, Moore SM, Ruland CM. How user characteristics affect use patterns in web-based illness management support for patients with breast and prostate cancer. J Med Internet Res 2013;15(3):e34 [FREE Full text] [doi: 10.2196/jmir.2285] [Medline: 23454601]
48. Hinz A, Krauss O, Hauss JP, Höckel M, Kortmann RD, Stolzenburg JU, et al. Anxiety and depression in cancer patients compared with the general population. Eur J Cancer Care (Engl) 2010 Jul;19(4):522-529. [doi: 10.1111/j.1365-2354.2009.01088.x] [Medline: 20030697]
49. Burgess C, Cornelius V, Love S, Graham J, Richards M, Ramirez A. Depression and anxiety in women with early breast cancer: five year observational cohort study. BMJ 2005 Mar 26;330(7493):702 [FREE Full text] [doi: 10.1136/bmj.38343.670868.D3] [Medline: 15695497]
50. Varsi C, Gammon D, Wibe T, Ruland CM. Patients' reported reasons for non-use of an internet-based patient-provider communication service: qualitative interview study. J Med Internet Res 2013;15(11):e246 [FREE Full text] [doi: 10.2196/jmir.2683] [Medline: 24220233]
51. 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]
52. Lluch M. Healthcare professionals' organisational barriers to health information technologies-a literature review. Int J Med Inform 2011 Dec;80(12):849-862. [doi: 10.1016/j.ijmedinf.2011.09.005] [Medline: 22000677]
53. Mair FS, May C, O'Donnell C, Finch T, Sullivan F, Murray E. Factors that promote or inhibit the implementation of e-health systems: an explanatory systematic review. Bull World Health Organ 2012 May 1;90(5):357-364 [FREE Full text] [doi: 10.2471/BLT.11.099424] [Medline: 22589569]
54. André B, Inger Ringdal G, Loge JH, Rannestad T, Laerum H, Kaasa S. Experiences with the Implementation of Computerized Tools in Health Care Units: A Review Article. International Journal of Human-Computer Interaction 2008 Dec 08;24(8):753-775. [doi: 10.1080/10447310802205768]
55. Wakefield DS, Mehr D, Keplinger L, Canfield S, Gopidi R, Wakefield BJ, et al. Issues and questions to consider in implementing secure electronic patient-provider web portal communications systems. Int J Med Inform 2010 Jul;79(7):469-477. [doi: 10.1016/j.ijmedinf.2010.04.005] [Medline: 20472495]

XSL•FO
RenderX

56. Nijland N, van Gemert-Pijnen J, Boer H, Steehouder MF, Seydel ER. Evaluation of internet-based technology for supporting self-care: problems encountered by patients and caregivers when using self-care applications. J Med Internet Res 2008;10(2):e13 [FREE Full text] [doi: 10.2196/jmir.957] [Medline: 18487137]
57. Singh H, Fox SA, Petersen NJ, Shethia A, Street RL. Older patients' enthusiasm to use electronic mail to communicate with their physicians: cross-sectional survey. J Med Internet Res 2009;11(2):e18 [FREE Full text] [doi: 10.2196/jmir.1143] [Medline: 19632972]
58. Curry SJ. eHealth research and healthcare delivery beyond intervention effectiveness. Am J Prev Med 2007 May;32(5 Suppl):S127-S130. [doi: 10.1016/j.amepre.2007.01.026] [Medline: 17466817]
59. Patel VL, Cytryn KN, Shortliffe EH, Safran C. The collaborative health care team: the role of individual and group expertise. Teach Learn Med 2000;12(3):117-132. [doi: 10.1207/S15328015TLM1203_2] [Medline: 11228898]
60. Grimsbø G, Ruland CM, Finset A. Cancer patients' expressions of emotional cues and concerns and oncology nurses' responses, in an online patient-nurse communication service. Patient Educ Couns 2012 Jul;88(1):36-43. [doi: 10.1016/j.pec.2012.01.007] [Medline: 22326865]

Abbreviations

CBI: Cancer Behavioral Inventory
IPPC: Internet-Based Patient Provider Communication service
HADS: Hospital Anxiety and Depression Scale
LLM: linear mixed model
MSAS: Memorial Symptom Assessment Scale
RCT: randomized controlled trial
SCQ-19: Self-Administered Comorbidity Questionnaire

Edited by G Eysenbach; submitted 13.03.14; peer-reviewed by T Baker, A Batenburg, L Donkin; comments to author 18.07.14; revised version received 29.08.14; accepted 08.11.14; published 18.12.14.

Please cite as:
Børøsund E, Cvancarova M, Moore SM, Ekstedt M, Ruland CM
Comparing Effects in Regular Practice of E-Communication and Web-Based Self-Management Support Among Breast Cancer Patients: Preliminary Results From a Randomized Controlled Trial
J Med Internet Res 2014;16(12):e295
URL: http://www.jmir.org/2014/12/e295/
doi:10.2196/jmir.3348
PMID:25525672

©Elin Børøsund, Milada Cvancarova, Shirley M Moore, Mirjam Ekstedt, Cornelia M Ruland. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 18.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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

Retaining Critical Therapeutic Elements of Behavioral Interventions Translated For Delivery via the Internet: Recommendations and an Example Using Pain Coping Skills Training

Christine Rini[1,2], PhD; Laura S Porter[3], PhD; Tamara J Somers[3], PhD; Daphne C McKee[3], PhD; Francis J Keefe[3], PhD

[1]University of North Carolina at Chapel Hill, Gillings School of Global Public Health, Department of Health Behavior, Chapel Hill, NC, United States

[2]UNC Lineberger Comprehensive Cancer Center, Chapel Hill, NC, United States

[3]Pain Prevention and Treatment Research Program, Department of Psychiatry & Behavioral Sciences, Duke University Medical Center, Durham, NC, United States

Corresponding Author:
Christine Rini, PhD
UNC Lineberger Comprehensive Cancer Center
School of Medicine, CB# 7295
Chapel Hill, NC, 27599-7295
United States
Phone: 1 919 966 3036
Fax: 1 919 966 3015
Email: christine.rini@unc.edu

Abstract

Evidence supporting the efficacy of behavioral interventions based on principles of cognitive behavioral therapies has spurred interest in translating these interventions for delivery via the Internet. However, the benefits of this dissemination method cannot be realized unless the translated interventions are as effective as possible. We describe a challenge that must be overcome to ensure this occurs—Internet interventions must retain therapeutic components and processes underlying the success of face-to-face interventions on which they are based. These components and processes vary in the ease with which they can be translated to the online environment. Moreover, some are subtle and may be overlooked, despite being recognized as essential to the success of face-to-face interventions. We provide preliminary guidance for retaining critical therapeutic components and processes in the translation process, using Pain Coping Skills Training for osteoarthritis pain to illustrate methods. Directions for future research are also discussed.

(J Med Internet Res 2014;16(12):e245) doi:10.2196/jmir.3374

KEYWORDS

psychotherapeutic processes; cognitive behavioral therapy; Internet; eHealth; intervention; treatment efficacy; musculoskeletal pain; osteoarthritis

Introduction

Interest in developing and disseminating effective behavioral interventions is driven by the critical role behavior plays in the most prevalent physical and mental health problems in the United States and elsewhere [1,2]. Empirical support for the efficacy of behavioral interventions has been shown in research focused on disease prevention (eg, seeking to increase health promoting behaviors or to decrease health-damaging behaviors [3-6]) as well as research focused on helping people manage symptoms of diagnosed physical and mental health problems (eg, persistent pain or depression [7-10]). For example, studies of behavioral approaches to treating persistent musculoskeletal pain, cancer pain, and other types of persistent pain are yielding growing evidence for the efficacy of interventions based on cognitive behavioral therapies (CBT; eg, [9,11-13])—the most commonly researched and used psychotherapeutic approach [10]. Increasing awareness of the individual and public health risks associated with pharmacologic treatments for persistent pain [14,15] may add impetus to efforts to develop evidence-based behavioral pain treatments.

Despite the growing number of empirically supported behavioral interventions, only a small proportion of people who could benefit from them are currently able to access them. Barriers that limit access to these interventions are often interrelated and

XSL•FO
RenderX

include difficulties integrating them into clinical care (eg, due to reimbursement problems or misconceptions held by health care providers about their efficacy or ideal use), the lack of available therapists in some geographic regions, the substantial time and resources needed to deliver them, the shortage of qualified therapists, and the need for people to travel to in-person sessions held at a set time [16,17].

As noted by a number of researchers [18-20], these and other barriers can be addressed by translating effective interventions for dissemination via the Internet, using interactive technologies that allow people to access them through home computers and mobile computing platforms. Internet interventions address the needs of people unable or unwilling to attend in-person treatment sessions (eg, because they are too busy, live in a remote area, lack transportation or other resources, have physical disabilities, dislike individual or group therapies, or are concerned about privacy or stigma). These individuals can complete these interventions at home, at their own pace, with access unlimited by time or day, reviewing information as often as they like. By eliminating the need for in-person meetings, Internet interventions also address barriers related to the substantial time and resources needed to deliver interventions in a face-to-face setting.

However, fully realizing these benefits requires that both interventions be efficacious. Evidence shows that Internet interventions can produce effect sizes comparable to effects of face-to-face interventions [21-25] or other forms of treatment, including medications [26]. Yet, their effects can also be disappointing even in studies that are relatively strong methodologically (eg, [27-29]). Small effect sizes suggest room for improvement, as do generally high attrition rates. One fundamental problem that may contribute to insufficient efficacy involves the extent to which Internet interventions maintain the integrity of the face-to-face interventions on which they are based. Some key components and processes of CBT-based therapies have been widely studied (eg, ways to increase self-efficacy for engaging in a behavior [30-32]). Others are less well understood (eg, subtleties in the ways that well-trained therapists interact with participants to increase engagement in intervention activities). Consequently, some of these components and processes will be more difficult than others to reproduce online. The goals of this paper are to highlight the importance of ensuring that critical therapeutic components are present and therapeutic processes are engaged when translating face-to-face interventions for Internet delivery, to provide preliminary guidance for doing so, and to describe directions for future research to support this effort.

CBT-based interventions share the assumption that people's behavioral and emotional responses are determined not solely by objective features of the world (eg, lack of reinforcement for remaining active), but rather by how people construe them; that is, their cognitive appraisals. Therapeutic components and processes of these interventions and typical methods through which they are delivered in face-to-face interventions arise from this assumption. Accordingly, these interventions typically teach participants to use skills for identifying, reality testing, and changing maladaptive cognitions (eg, beliefs, expectations, attributions) and behaviors (eg, a sedentary lifestyle, fearful avoidance of activities they are capable of doing) that contribute to current physical or psychological symptoms or increase risk for future symptoms [33,34].

In this discussion, we will use Pain Coping Skills Training (PCST) for osteoarthritis (OA) pain [35] as a case study to illustrate methods for translating a face-to-face CBT-based intervention for delivery as an Internet-based program. PCST is most often presented by a trained therapist in an individual or group setting over the course of 10-12 weeks. The skills taught in PCST have been shown to reduce pain and disability caused by OA, cancer, and other disorders [36-41]. They target cognitions involving pain catastrophizing (exaggerated negative beliefs about pain [42]) and behaviors involving learned, maladaptive responses to pain such as spending excessive time reclining. For example, people with persistent pain engaging in PCST learn how to pace physical activities so as to gradually increase their activity and return to activities they found meaningful and enjoyable (called "activity pacing" or "activity-rest cycling" [43,44]).

Translating Face-to-Face Cognitive Behavioral Interventions for Internet Delivery

Overview

This section summarizes important therapeutic components and processes in CBT-based interventions, as they are applied in PCST, describes methods used to engage them in face-to-face PCST, and explains our methods for maintaining their integrity when translating PCST to an Internet intervention called PainCOACH. Rather than providing an exhaustive list of all possible methods, our intent is to offer a case study and to encourage research to guide development of best practices for future translation efforts. We begin by summarizing the structure of the PainCOACH program and our general translation approach and then discuss decisions we made about how to implement specific therapeutic components and processes of face-to-face PCST in the online environment.

PainCOACH Structure

PainCOACH is an automated program (ie, it does not include therapist contact) that includes eight 35- to 45-minute training sessions, each teaching a cognitive or behavioral coping skill drawn from face-to-face PCST. Participants complete one session per week over 8 weeks. Session 1 starts with an overview of the PainCOACH program, PCST, and the intervention's therapeutic rationale (a simplified version of gate control theory) [45]. This overview is followed by training in the first pain coping skill: progressive muscle relaxation. Sessions 2-7 teach, respectively, brief relaxation skills (ie, "mini-practices"), activity-rest cycling, pleasant activity scheduling, cognitive restructuring ("coping thoughts"), pleasant imagery, and problem solving. The eighth session consolidates learning and teaches strategies for long-term skill use. Between sessions, participants are asked to practice their newly learned skill and any skills they learned in past sessions.

XSL•FO
RenderX

The program also includes supplementary modules to enhance engagement and facilitate practice. One of these modules (COACHtrack) allows participants to self-monitor their progress by reviewing and changing practice goals, recording practices and "coping confidence" (self-efficacy), viewing graphic summaries of progress over time, and managing automated practice reminders. A second module (COACHchat) allows participants to read about others' experiences using the pain coping skills and to submit descriptions of their own experiences. COACHchat was designed to provide some of the benefits of observational learning [46] that people get in face-to-face PCST when it is provided in small groups. Finally, the program includes a module (MyCOACH) that provides information about the program, the study team, and actions to take in a medical or mental health emergency.

The PainCOACH home page (Figure 1) provides a portal to these modules, the training sessions, and other program features. One additional feature on the home page is a "toolbox" icon that appears next to each session once it has been completed. Selecting the toolbox gives participants easy access to segments of completed training sessions they might like to review. The home page also displays reminders and encouraging messages as well as badges participants earn by completing training sessions and selected tasks (eg, practicing skills, or reading or posting experiences in COACHchat).

The modules, sessions, and features are controlled by a programming component that applies an "expert systems" approach [47]. That is, it pairs decision rules (in the form of tailoring algorithms) with a knowledge database to simulate the behavior and judgment of an expert—in this case, a highly trained therapist experienced in delivering face-to-face PCST. The decision rules customize (tailor) participants' experience in the program based on their responses and progress through the program. Examples of tailoring are provided in the next section.

The program was developed using standard usability guidelines [48], to ensure that it would be easy to use and appropriate for a diverse population (eg, consistency in screen design and navigation, easy-to-read text, understandable language). Usability was also enhanced by ensuring that each of PainCOACH's eight sessions share a similar structure. Briefly, sessions begin with a review of practice completed in the prior week (except in session 1, where this segment is replaced by the introduction). Next, a new skill is introduced and participants are given the chance to try it and to evaluate their experience. After tailored feedback and remediation are provided, participants are reminded of the importance of practicing learned skills. They are taught strategies for prompting the practice of each new skill in the coming week. Sessions end with a reminder to complete the subsequent session in a week and to use COACHtrack and COACHchat.

Figure 1. PainCOACH Home Page.



Translation Approach

Our challenge was to translate a highly interactive, collaborative face-to-face intervention to an automated Internet intervention, using technology to detect and address problems that would normally be handled by a trained therapist. For instance, in a face-to-face setting a therapist can monitor participants' verbal and non-verbal cues for signs that they do not understand or agree with information being communicated—problems that would reduce the efficacy of the intervention. Our task was facilitated by the fact that we began with a well-characterized intervention. In addition, we relied on specialized clinical expertise to ensure that activities participants completed in PainCOACH were as likely as possible to have the same therapeutic effects as activities they would complete in face-to-face PCST; our development team included 4 clinical psychologists with over 90 years of combined experience delivering face-to-face PCST and similar therapies—our expert therapists. Discussion of our translation process highlights some of the critical guidance they provided.

Translation began with screen-by-screen plans laying out the PainCOACH program's content and function, based on content from materials and scripts used in face-to-face PCST sessions. To ensure that participants would not only learn about skills, but also master how to use them, these initial plans were guided by principles of Social Cognitive Theory [46,49], adult learning [50,51], and multimedia instruction [52]. For instance, this body of literature led us to include simple interactive exercises to enhance mastery of new skills, persuasive arguments regarding participants' ability to engage in target behaviors (social persuasion), and social modeling. With respect to the latter, we supplemented educational content with information presented by "characters"; that is, actors (an African American man, a

XSL•FO
RenderX

Latina woman, and a white woman) selected to represent average members of our population in terms of their physical appearance and age. Each was given a name and persona that carried throughout the program. They reinforced main points and modeled use of skills in daily life, including barriers to using skills and ways to overcome those barriers.

The development team then met to refine the initial plans, discussing intervention components or processes that posed special challenges for translation. A primary focus of these discussions was to develop solutions that ensured the program would engage therapeutic components and processes present in face-to-face PCST, as shown in Table 1, column A. In most cases our solutions represented our expert therapists' clinical judgment because there is currently little empirical evidence to guide these kinds of decisions. We incorporated our expert therapists' knowledge and experience into the PainCOACH program through its content libraries and "expert systems" tailoring algorithms, described above. At times, we tried to ensure the user experience would be consistent with features of face-to-face PCST (eg, in terms of the information and feedback provided, how it was worded, training techniques, and decision rules) [47]. At other times more adaptation was needed. We describe informative examples in the next section.

Translating Intervention Content

As shown in Table 1, column A, the first therapeutic component of PCST (and other CBT-based interventions) involves providing participants with relevant knowledge. In PCST, this knowledge begins with a plausible therapeutic rationale [53] explaining the causal link between use of pain coping skills and the ability to attain positive, desired outcomes (eg, reduced pain). The rationale provides a framework for learning skills and promotes a sense of control and positive outcome expectancies, leading to better treatment outcomes [54,55]. For PCST, the rationale uses a simplified description of Melzack and Wall's [45] gate control theory to explain how behaviors, cognitions, and emotional responses can affect the experience of pain [53]. Additional knowledge provided in face-to-face PCST sessions focuses on explaining the skills, and providing written materials and graphics to illustrate concepts and reinforce learning.

One goal of translating this educational content to PainCOACH was to enrich it with multimedia capabilities of the online environment. Guided by usability guidelines as well as principles of adult learning [50,51] and multimedia instruction [52], animation and graphics were strategically incorporated to draw attention to key points, and simple interactive exercises were used to support learning and engagement. Participants controlled the pace of learning by using simple buttons (eg, "next" or "back") to advance through screens that presented "chunks" of related information. In addition, research on relational agents and participant input led us to attempt to increase engagement, accountability, and adherence by minimizing onscreen text and instead having a female virtual Coach [56,57] guide participants through the PainCOACH program. An audio track of the Coach's voice runs throughout the program. She greets participants at the beginning of each session and provides verbal instruction, feedback, and encouragement throughout the session. Her manner is conversational, empathetic, and warm. Pedagogical agents with these kinds of characteristics can promote better learning by cueing people to respond as if they are engaging in a social interaction, in turn promoting social bonding and greater motivation and cognitive processing [58-60]. All audio and onscreen text used simple, conversational language.

During the translation process, our team's expert therapists identified the concepts or skills participants were most likely to have concerns about. Rather than having concerns linger and interfere with learning, screens addressing common concerns are used in key places throughout the program. In some places they are presented on buttons participants can press to get more information, and the responses they get are presented using language trained therapists would use in face-to-face PCST. Alternatively, in some parts of the program we addressed common concerns by having our "characters" provide social persuasion and modeling. For instance, if our team's expert therapists indicated that some participants question the value of a particular skill, we might have a character appear onscreen talking about how he questioned its value but then tried it and found it to be helpful (eg, "I didn't think this would work for me, but I decided to give it a try…"). Of course, it is possible that some participants will have concerns we did not anticipate and address; this is a downside of an automated program. However, because we were guided by therapists with considerable experience delivering PCST, we were confident we had covered the most common concerns that could interfere with learning.

One skill, cognitive restructuring, involves teaching participants to recognize and challenge overly negative thoughts that affect their pain. It presented a challenge to translate to PainCOACH because it is relatively abstract and therefore potentially difficult for participants to understand and apply. We addressed this challenge, in part, by breaking education about this skill into two sessions. First we introduced the concept of "unhelpful automatic thoughts" with animated examples and social modeling to show examples in the context of daily life. Then participants were asked to use a worksheet (provided in a printed program workbook) to track their unhelpful automatic thoughts, the situations in which they occurred, and any emotional and physical consequences they noticed in the coming week. We reasoned that completing this homework and receiving feedback on these initial experiences would set the stage for the next session's training in how to reduce unhelpful automatic thoughts using "coping thoughts" (ie, cognitive restructuring), which included interactive exercises that helped participants apply this new educational material to experiences they had tracked in their prior week's homework.

It is important to note that education alone is not sufficient for ensuring mastery of pain coping skills, nor is it sufficient for ensuring that participants can use the skills appropriately and consistently in daily life. Rather, education is part of a more comprehensive training approach. One distinctive feature of PCST, as with CBT-based interventions in general, is that face-to-face training is delivered in sessions that are highly collaborative (the second therapeutic component in Table 1, column A). Therapists use a variety of methods to help

XSL•FO
RenderX

participants master new skills and to increase their motivation to use them regularly in daily life—a necessity if the skills are to have a positive influence on health and well-being.

Some methods therapists use to achieve these goals in face-to-face PCST are summarized in Table 1, column B, and examples of ways this component was implemented in PainCOACH are summarized in column C. For instance, therapists use behavioral rehearsal to help participants develop a first-hand understanding of how it feels to use a skill effectively. They typically model a new skill, guide participants as they complete experiential exercises (eg, role play, guided discovery, behavioral experiments), observe participants' performance to ensure adequate learning, and provide corrective feedback and encouragement. These exercises help participants build proficiency and self-efficacy for using the skills [49] and reinforce positive outcome expectancies, in turn increasing the probability they will have positive experiences that increase their motivation to use the skills.

One example of how collaborative skills training was translated to PainCOACH is a process we referred to as "unpacking" a participant's experience with using a new skill. Unpacking is a process through which participants learn to be aware of their emotional, cognitive, and behavioral responses to practicing a skill—how and where they are doing it, how it makes them feel, which aspects of the practice worked well or poorly, and how they might alter their practice to be more effective in the future. Its goal is to help them understand and benefit from practice. In a face-to-face PCST, the therapist would ask a series of questions to highlight features of the practice experience (eg, "what changes did you notice in your body or in what you were feeling or thinking while you practiced?"), monitoring participants' verbal and nonverbal responses to address unspoken thoughts, concerns, or questions. These activities help participants recognize and appreciate positive experiences during practice. They also build participants' self-efficacy for recognizing and solving problems that lead to negative experiences that can reduce motivation to use skills. In sum, unpacking is complex and subtle yet critical for learning and, ultimately, behavior change. We believe this important process is likely to be overlooked in unguided Internet interventions and that it is likely to be used insufficiently in guided Internet interventions that include minimal therapist contact [61], or when the therapist role is enacted by individuals who are not trained therapists.

For PainCOACH, we developed an interactive unpacking exercise that participants completed after practicing selected skills. For instance, participants were educated about progressive muscle relaxation in Session 1, then followed the Coach's audio instructions to try the skill. After they finished the practice, the Coach walked them through the unpacking exercise. The first step in the unpacking exercise was designed to address a challenge that the team's expert therapists alerted us to—participants do not always notice or pay attention to sensations and other experiences they have when they first practice a new skill. This kind of self-examination is new and unfamiliar to most participants. Consequently, we chose to have our three characters model this process. Specifically, the Coach introduced the idea that people have different experiences when they try progressive muscle relaxation and then explained to participants that they could hear the characters describe their first experience with progressive muscle relaxation by selecting each character's picture. The characters' responses were designed to model how a person would notice and describe common positive or negative experiences with progressive muscle relaxation: William mentioned a common physical sensation (legs feeling heavy, like they were sinking into his chair), Linda noted a sense of calm that lasted even after the practice ended, and Patricia discussed being distracted by her thoughts before realizing she could put her worries aside for a while to focus on relaxing (modeling a solution).

The next step in the unpacking exercise was to allow participants to report their own experiences and receive feedback from the Coach. Continuing the session without having the Coach ask about and respond appropriately to participants' experiences would frustrate those who had negative experiences (eg, because the Coach did not understand their problems and demonstrate concern about their learning) while missing an important opportunity to reinforce positive experiences. A responsive "dialogue" between the Coach and participants was important for building the working alliance so critical to CBT-based therapies. In order to implement this dialogue, PainCOACH's expert system needed information that could be used to tailor the Coach's response. We chose a simple interactive exercise to get this information. The Coach asked participants what they had experienced during their practice, and they responded by dragging and dropping boxes labeled "more of this", "less of this", and "no change in this" into boxes labeled with positive and negative experiences commonly reported by participants in face-to-face PCST. Thus, a participant might report a relatively positive experience by reporting increases in feeling relaxed, happy, and clear-headed and decreases in pain, stiffness, and muscle tension, with no change in other phenomena. Another participant might report a relatively negative experience, such as increases in feeling fidgety and sleepy, with no decrease in muscle tension. Participants' responses were used by the tailoring algorithm to provide appropriate feedback (eg, reinforcing positive experiences or normalizing and addressing negative experiences). Finally, the unpacking exercise ended with the Coach introducing a list of common concerns and problems that participants could select to get reassurance, normalization, and tips on how to resolve them. The goal was to resolve that could reduce participants' motivation or self-efficacy for using the skill. All of these steps in the PainCOACH unpacking exercise were evaluated in user testing, as described below, to ensure they were intuitive to use and made participants feel understood and supported.

The third therapeutic component of PCST (shown in Table 1, Column A) involves self-monitoring. Participants are taught to recognize pain-relevant patterns of behavior, emotions, and cognitions in their daily life. In face-to-face PCST, therapists provide worksheets to facilitate self-monitoring and discussion, in turn helping participants recognize when and how to use pain coping skills (see column B). As summarized in column C, we implemented self-monitoring in PainCOACH by having the Coach teach participants to complete worksheets (provided in a printed program workbook), illustrating this task with animated

XSL•FO
RenderX

graphics, then having participants complete interactive exercises to learn how to recognize their own maladaptive behaviors, emotions, and cognitions and the situations in which they occur.

The next therapeutic component of PCST (the fourth component in Table 1, Column A) helps ensure that participants are able to use skills appropriately and stay motivated to do so over time. This is a critical part of training because health benefits require that the skills are used in daily life. Examples of ways therapists deliver this component in face-to-face PCST are summarized in column A. For instance, participants are expected to practice using the skills outside of sessions to ensure they gain experience using skills in various life situations, develop confidence in their ability to use them, and develop an ability to apply them in a flexible manner to meet varying challenges of daily life. In PCST, like other CBT-based pain interventions, participants are not expected to use all of the skills they learn all of the time. Rather, they learn to identify the skills they find most helpful and to incorporate them into a plan to prevent or manage pain flares.

It is also important to understand participants' experiences with practice; they are not likely to keep using skills if they do not find them to be useful or if they dislike using them for other reasons. Accordingly, after participants learn a new skill in face-to-face PCST, they work with their therapist to develop goals and plans for practice. Then, in their next session, therapists and participants review whether practice sessions and goals were completed and discuss successes and/or problems participants experienced. The therapist helps participants recognize successful experiences, reinforcing them and problem solving as necessary.

Features of the PainCOACH program summarized in Table 1, column B, were designed to correspond with these important activities. The Coach explains the importance of practice early in the program. That message is reiterated throughout the program, and various program features support and reward practice. For instance, at the end of each training session the Coach suggests strategies for remembering to practice, then leads participants through an interactive exercise to set practice goals for the newly learned skill. Participants are periodically reminded to update their goals in COACHtrack, which also allows them to record practices they complete between sessions and to generate a graph showing their progress over time. They are rewarded for completing practices with badges recognizing their accomplishments. In addition, at the beginning of each session, the Coach reviews participants' completion of practices in the prior week. She uses any practices participants entered in COACHtrack as a starting point (eg, "I see from COACHtrack that you practiced your activity/rest cycle last week…") or allows them to indicate a number before continuing. For participants who have not practiced, the Coach provides information to troubleshoot reasons this may have occurred and reiterates the importance of practice, guided by tailoring algorithms to ensure the information and suggestions are appropriate. The wording of these communications was carefully crafted to reflect approaches therapists use in face-to-face training (eg, to develop a strong working alliance, as described next). PainCOACH also lets participants set up automated practice reminders delivered as an email, text message, or telephone message.

Session 8 of the program also promotes ongoing use of the skills through its focus on preparing participants to continue using skills over time. Participants complete exercises that involve thinking about which skills they have used most and found to be most helpful. They also play a game that asks them to consider the kinds of situations that might lead them to use different skills. It presents scenarios, asks them to select which skills they would use in the scenarios, and provides feedback on their selection of skills. Other exercises help them think about setbacks that might reduce their use of skills and how to overcome them.

The last key component of CBT-based interventions shown in Table 1 [45,50,62], column A, is the working alliance that therapists develop with participants. Specifically (as summarized in column B), they provide treatment in a warm, accepting manner that communicates partnership in the process of behavior change. They assess and monitor participants' ongoing level of understanding, motivation, and investment in the intervention and the factors that interfere with treatment success. They also adjust the intervention as needed, based on their knowledge of psychosocial factors that affect participants' engagement in treatment. Taken together, these methods enable the development of a working alliance.

It is currently unclear whether it is critical to develop this kind of alliance in Internet interventions. Some research suggests that it might not be [63,64], but other research has found that developing a greater alliance is associated with better outcomes in Internet interventions [65]. More research is needed to clarify whether the importance and nature of the working alliance differs in face-to-face versus Internet interventions. When translating face-to-face PSCT for delivery in PainCOACH, we considered it to be important and took steps to ensure participants developed the sense that behind the program text, audio, and exercises was an empathetic therapist who understood them, cared about their pain, and was collaborating with them to find ways to manage it (summarized in column C). For instance, our decision to have the program led by a virtual Coach was motivated in large part by a desire to facilitate development of a strong working alliance. As noted above, this program feature was also expected to increase engagement, accountability, and adherence [53,54]. Images of the Coach (which change so that her body language and facial features correspond to ideas being expressed) appear onscreen frequently. Her manner is conversational, empathetic, and warm. She greets participants when they begin sessions and speaks directly to them throughout the program, including setting up an appointment with them to return for the next week's session. Tailoring algorithms are used to ensure her feedback and responses are personalized so that participants feel that they are understood. Her dialogue was carefully crafted to use techniques that therapists are trained to use to promote development of a working alliance in face-to-face PCST (and other CBT-based interventions).

XSL•FO
RenderX

Table 1. Therapeutic components and processes of cognitive behavioral therapies.

A	B	C
Therapeutic components of CBT interventions	Methods for delivering them in face-to-face interventions	Examples of methods for delivering them in Internet-based interventions
1. Relevant knowledge (education) Therapeutic rationale: Simplified explanation of Gate Control Theory [45] and how cognitions, behaviors, and emotions and description of how changing them would achieve treatment goals. Information about skill: Description of skill, referencing processes described in rationale, along with explanation of how and when it is used in daily life.	Therapist provides verbal education in highly interactive and collaborative manner	Multimedia education; Interactive exercises to assess learning and provide remediation
		Concerns addressed through text covering "frequently asked questions", social modeling (ie, stories told by characters similar to the target audience), or providing ability to ask questions and get personalized response (eg, via email)
		Ability to revisit material as needed
	Therapist provides written educational materials to reinforce learning	Workbook or ability to print computer-tailored materials
	Graphics used to illustrate concepts and reinforce learning	Graphics enhanced with audio, animation, and testimonials (to provide social modeling)
2. Collaborative skills training, focused on developing mastery of skills and self-efficacy for using them in daily life	Therapist models skill to demonstrate its proper use	Skill modeled in video or with graphics showing therapist or simulated participant (for social modeling)
		Ability to revisit material as needed
		Easy to access supportive resources (eg, audio recordings to guide practices available for download or accessible through mobile technologies)
	Therapist guides participant through coping skills practice (behavioral rehearsal)	Program guides participant through behavioral rehearsal
	Therapist observes performance to ensure adequate learning, providing corrective feedback	Interactive exercises help participant identify problems experienced during behavioral rehearsal
		Computer-tailored corrective feedback
	Therapist solicits information about participant's experience with behavioral rehearsal	Interactive exercises to gather information about participant's experience with behavioral rehearsal
	Therapist highlights positive experiences	Interactive exercises to help participant recognize positive experiences
	Therapist provides feedback about performance to promote self-efficacy and positive outcome expectancies	Computer-tailored feedback appropriate to participant's report of problems and positive experiences during behavioral rehearsal
		Interactive exercise to assess and promote self-efficacy and positive outcome expectancies
	Therapist instructs participant in how to use skills and associated worksheets outside of sessions	Multi-media instruction in how to use skills and worksheets outside of sessions
		Examples of skill and worksheet use modeled by simulated participants
		Interactive exercises to help participant develop personalized plan describing where and when skills and worksheets will be used
	Participant practices skill outside of session	Participant sets up and receives automated practice reminders
		Participant practices skill outside of session
	Participant tracks practices and experiences with paper worksheet	Participant tracks practices and experiences with electronic application (eg, on mobile device)
		Immediate computer-tailored feedback to reinforce practice
		Graphics show progress over time

XSL•FO
RenderX

A	B	C
Therapeutic components of CBT interventions	Methods for delivering them in face-to-face interventions	Examples of methods for delivering them in Internet-based interventions
	At subsequent session, therapist reviews worksheet with participant to provide feedback regarding frequency and adequacy of skill use	At subsequent session, interactive exercises elicit number of practices completed and experiences during practices
		Interactive exercises identify and troubleshoot problems and draw attention to positive experiences
	Therapist provides feedback to reward skill practice and to reinforce future practice	Immediate computer-tailored feedback about frequency and adequacy of skill use
		Badges, unlocked features, or other rewards to recognize skill practice and to reinforce future practice
3. Self-monitoring: Recognizing and recording experiences to promote use of skills	Participant and therapist identify participant's maladaptive cognitions, behaviors, and emotions and the internal and external cues associated with them (eg, emotions, environmental factors, social/interpersonal factors).	Interactive exercises help participant identify maladaptive cognitions, behaviors, and emotions and the internal and external cues associated with them (eg, emotions, environmental factors, social/interpersonal factors).
	Therapist shows participants how to complete worksheets that facilitate self-monitoring (eg, "thought records" used to record thoughts associated with negative responses)	Multimedia instruction in how to complete worksheets
		Electronic worksheets (eg, using mobile technologies)
4. Ensure appropriate, enduring use of skills in daily life	Therapist works with participant to set appropriate practice goals	Multimedia education on setting appropriate practice goals
		Interactive exercises help participant set goals
	Therapist works with participant to develop personalized strategies for meeting goals	Multimedia education showing strategies for meeting goals (including modeling by simulated participants)
		Ability to set up automated reminders (eg, phone calls, email, text messages, coordination with wearable monitors)
	Therapist provides worksheet to record goals and track goal-related behaviors	Participant records goal-related behaviors in program (eg, on computer or mobile device)
		Graphical display show goal attainment over time
	Therapist works with participant to identify barriers to goal attainment and problem solve as necessary	Interactive exercises to identify barriers to goal attainment, with computer-tailored guidance for how to overcome them
	Therapist works with participant to increase goals over time, being responsive to participant preferences	Participant periodically reminded to update goals
		Program assists with setting new goals using computer-tailoring based on practice history and goal attainment
	Therapist recognizes goal attainment to enhance participant self-efficacy	Badges, unlocked features, or other rewards to recognize goal attainment
	Discussion with therapist regarding possible future setbacks that could affect ongoing use of skills	Multimedia education describing possible future setbacks, modeled by simulated participants
		Interactive exercise to generate personalized list of possible future setbacks that could affect ongoing use of skills
	Therapist guides participant in developing a plan to cope with setbacks	Interactive exercise to develop plan to cope with setbacks (eg, applying research on implementation intentions to create a specific if-then plan) [62]
5. Therapeutic/ working alliance	Therapist behaviors (warmth, acceptance, collaborative approach) promote therapist-participant bond and ability to agree on goals and tasks	Use of "virtual coach" (a relational agent; [57]) to build social-emotional relationships with participant by emulating face-to-face interaction with therapist

XSL•FO
RenderX

A	B	C
Therapeutic components of CBT interventions	Methods for delivering them in face-to-face interventions	Examples of methods for delivering them in Internet-based interventions
	Therapist assesses participant's engagement and agreement with goals and methods at various points during therapy	Algorithms evaluating participant engagement through program use, practice, and goal attainment
		Assessments to measure participant's agreement with goals and methods at various points during therapy
		Computer-tailored messages to provide personalized feedback in a way that communicates warmth, acceptance, and collaborative approach

User and Pilot Testing

Our iterative development process included both user testing and pilot testing, with the overarching goal of ensuring the PainCOACH program was technologically sound, engaging, and easy to understand and use by our target population—older adults with OA of the hip or knee and OA-associated persistent pain, some of whom would be rural and of low socioeconomic status.

Two rounds of user testing were conducted early in the development process with 49 participants drawn from our target population, selected to be diverse in terms of gender (female: 55%, 27/49), race (Non-Hispanic white: 51%, 25/49; African American: 49%, 24/49), education (from less than high school to graduate degrees), and experience using computers and the Internet (from none to daily). They ranged in age from 53-87 years old (with an average age of 67). Groups of participants were provided with an introduction to the PainCOACH program and study goals, then met one-on-one with a trained study staff member who observed them as they worked through selected sessions and used a semistructured interview to gather feedback. In the first round, participants provided feedback on the content and functionality of a working prototype of the program's home screen and session 2. In the second round, participants provided feedback on the content and functionality of the revised session 2, plus two additional sessions. Across both rounds of user testing, positive feedback indicated a high degree of enthusiasm for learning PCST and using an Internet intervention for doing so. Many features of the program were well liked, including the overall look of the program, the Coach, learning about others' experiences using PCST (ie, the characters), the information that was provided, and the skills being taught. Feedback also revealed a need for various refinements, including redesigning the home screen to clarify its features and functions, eliminating most exercises that required typing (because of confusion about using the tablet computer's onscreen keyboard), reducing text and making it larger, and making some navigation features and exercises more intuitive.

When a working prototype of the entire program was complete, eight participants pilot tested it at home. All had OA of the hip or knee and were screened for having significant OA pain. The sample was 75% female (6/8), 63% Non-Hispanic white (5/8), and 38% African American (3/8), aged 54-76 (average age 68 years), and diverse in terms of annual income (less than US $15,000 to approximately $75,000), education (less than high school to a graduate degree), and experience using computers and the Internet (none to daily). After an in-person orientation meeting, participants were asked to complete the PainCOACH program within 2 weeks. Our goal was not to have them learn and use the pain coping skills, but rather to provide feedback on technical problems and any aspects of the program that were difficult, confusing, or frustrating to use. Pilot test participants were given a notebook with screenshots to aid in noting and reporting problems, which were communicated to us throughout pilot testing in brief phone interviews. At the conclusion of the pilot test, participants completed a questionnaire to report usability. Overall, feedback was highly positive. For instance, most participants strongly agreed or agreed that it was easy to log into PainCOACH and navigate through it, they felt confident using it, and they thought it was easy to use. All agreed or strongly agreed that the information was useful, that it made them feel that pain coping skills were important to learn, and that the exercises and stories about other people's experiences helped them understand the information. The pilot test guided final refinements to the program, with most changes involving needing to optimize the speed at which the program's audio downloaded and to ensure participants had a strong enough Internet connection to access the program. Once these final refinements were made, we began a small-scale randomized controlled trial to evaluate its effects on key participant outcomes, acceptability, and feasibility. Feedback from the trial, in addition to evidence for the program's efficacy, will be used to make additional refinements for a larger scale evaluation.

Discussion

The benefits of delivering proven behavioral interventions via the Internet are compelling and address goals related to individual and public health. However, achieving these goals requires that the Internet versions of these interventions be as effective as possible. In the present paper, we highlighted one critical, but underappreciated challenge that must be overcome to ensure this occurs—behavioral Internet interventions must retain therapeutic components and processes underlying the benefits of the face-to-face interventions on which they are based (eg, knowledge, collaborative skills training, self-monitoring, ensuring appropriate and enduring use of skills, and working alliance). In the case of CBT-based interventions, these components and processes vary in the ease with which they can be translated to an interactive, online environment. Moreover, some are subtle and may be overlooked in the

XSL•FO
RenderX

translation process. Failing to engage them is likely to limit the efficacy of otherwise promising interventions, including those that are automated and therefore do not include therapist contact.

Unfortunately, little empirical evidence exists to guide decisions that must be made during the translation process. In light of the dearth of evidence, we argue that these interventions are more likely to be maximally effective if they are developed in close collaboration with expert therapists. Their professional training in the clinical application of the science of behavior change, paired with their extensive hands-on experience delivering an intervention in person, allows a development team to anticipate parts of the intervention during which participants will have problems learning and applying skills; the specific problems they will have; the best time-tested solutions for resolving the problems; and how to deliver those solutions so that participants are most likely to accept and benefit from them. Expert therapists are uniquely able to provide this critical input.

Close collaboration with expert therapists is likely to be especially critical for automated interventions that do not include any interaction with a therapist because these interventions depend solely on the adequacy of the computer program for their benefits. That is, their involvement facilitates use of an "expert systems" approach [47], which we implemented by approximating a real clinical encounter (eg, how a therapist assesses participants' needs, makes decisions based on those needs, and provides feedback in face-to-face PCST). We have highlighted one approach, but others certainly exist and more will arise as new technologies emerge and mature. Of course, each approach must be evaluated with research. It seems likely that research could find other effective strategies for engaging therapeutic processes online. Theoretical models of behavior change developed specifically for Internet interventions may be useful as a starting point for this work [66,67], as can research showing the most effective components of behavioral interventions [68].

Indeed, it has been noted that simulating interactions as they occur in a face-to-face setting may not always be the ideal goal. Rosser et al [69] propose that differences in how people interact with other people compared to how they interact with machines—or with other people via machines—will sometimes require therapeutic methods to be adapted rather than directly translated for Internet delivery. We agree and add that methods of engaging therapeutic processes may also be fundamentally changed—and potentially strengthened—by technological capabilities of the Internet. Consider, for instance, the goal of educating participants. Compared to written materials and graphics used in the face-to-face setting, it may be possible to enhance learning and mastery of new information through use of multimedia skills training paired with experiential exercises in which participants apply new knowledge to solve problems [52,70]. Likewise, new sensor technologies and smartphone applications may enable more effective self-monitoring than paper-and-pencil worksheets. Sophisticated use of emerging technologies such as wearable sensors, natural language processing, and machine learning could make self-monitoring less burdensome and more likely to help people detect and change maladaptive behaviors. Best practices for these and other uses of technology will need to be established with research.

If PainCOACH succeeds in improving pain and related outcomes, its success will provide only indirect evidence for the importance of ensuring that CBT-based Internet interventions (and particularly automated interventions) maintain the integrity of their corresponding face-to-face interventions. The trial testing PainCOACH was not specifically designed to evaluate the program's engagement of the same therapeutic processes as those engaged by face-to-face PCST. Conversely, if PainCOACH does not succeed in improving participant outcomes, it will not necessarily invalidate our argument; there are many reasons an Internet intervention may fail. Thus, the strongest foundation for development of CBT-based Internet interventions requires research focused specifically on evaluating the importance of each therapeutic component and establishing best practices for translating them to the Internet. Similar to usability standards that help researchers develop Internet-based interventions that are easy to use and engaging, it should be possible to develop empirically supported standards for translating or adapting critical therapeutic components and processes of face-to-face CBT and other therapies to the Internet. Innovation will always be important, but these kinds of standards would be an invaluable starting point to ensuring more efficient use of resources.

Acknowledgments

Research reported in this publication was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases, part of the National Institutes of Health, under Award Number R01AR057346. 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. Glanz K, Bishop DB. The role of behavioral science theory in development and implementation of public health interventions. Annu Rev Public Health 2010;31:399-418. [doi: 10.1146/annurev.publhealth.012809.103604] [Medline: 20070207]
2. McDermott MS, While AE. Maximizing the healthcare environment: a systematic review exploring the potential of computer technology to promote self-management of chronic illness in healthcare settings. Patient Educ Couns 2013 Jul;92(1):13-22. [doi: 10.1016/j.pec.2013.02.014] [Medline: 23566427]

XSL•FO
RenderX

3. Hobbs N, Godfrey A, Lara J, Errington L, Meyer TD, Rochester L, et al. Are behavioral interventions effective in increasing physical activity at 12 to 36 months in adults aged 55 to 70 years? A systematic review and meta-analysis. BMC Med 2013;11:75 [FREE Full text] [doi: 10.1186/1741-7015-11-75] [Medline: 23506544]
4. Appel LJ, Champagne CM, Harsha DW, Cooper LS, Obarzanek E, Elmer PJ, Writing Group of the PREMIER Collaborative Research Group. Effects of comprehensive lifestyle modification on blood pressure control: main results of the PREMIER clinical trial. JAMA 2003;289(16):2083-2093. [doi: 10.1001/jama.289.16.2083] [Medline: 12709466]
5. Turk MW, Yang K, Hravnak M, Sereika SM, Ewing LJ, Burke LE. Randomized clinical trials of weight loss maintenance: a review. J Cardiovasc Nurs 2009;24(1):58-80 [FREE Full text] [doi: 10.1097/01.JCN.0000317471.58048.32] [Medline: 19114803]
6. Mottillo S, Filion KB, Bélisle P, Joseph L, Gervais A, O'Loughlin J, et al. Behavioural interventions for smoking cessation: a meta-analysis of randomized controlled trials. Eur Heart J 2009 Mar;30(6):718-730 [FREE Full text] [doi: 10.1093/eurheartj/ehn552] [Medline: 19109354]
7. Golightly YM, Allen KD, Caine DJ. A comprehensive review of the effectiveness of different exercise programs for patients with osteoarthritis. Phys Sportsmed 2012 Nov;40(4):52-65 [FREE Full text] [doi: 10.3810/psm.2012.11.1988] [Medline: 23306415]
8. Barth J, Munder T, Gerger H, Nüesch E, Trelle S, Znoj H, et al. Comparative efficacy of seven psychotherapeutic interventions for patients with depression: a network meta-analysis. PLoS Med 2013;10(5):e1001454 [FREE Full text] [doi: 10.1371/journal.pmed.1001454] [Medline: 23723742]
9. Keefe FJ, Porter L, Somers T, Shelby R, Wren AV. Psychosocial interventions for managing pain in older adults: outcomes and clinical implications. Br J Anaesth 2013 Jul;111(1):89-94 [FREE Full text] [doi: 10.1093/bja/aet129] [Medline: 23794650]
10. Butler AC, Chapman JE, Forman EM, Beck AT. The empirical status of cognitive-behavioral therapy: a review of meta-analyses. Clin Psychol Rev 2006 Jan;26(1):17-31. [doi: 10.1016/j.cpr.2005.07.003] [Medline: 16199119]
11. Eccleston C, Morley SJ, Williams AC. Psychological approaches to chronic pain management: evidence and challenges. Br J Anaesth 2013 Jul;111(1):59-63 [FREE Full text] [doi: 10.1093/bja/aet207] [Medline: 23794646]
12. Eccleston C, Palermo T, Williams A, Lewandowski Holley A, Morley S, Fisher E, et al. Psychological therapies for the management of chronic and recurrent pain in children and adolescents. Cochrane Database Syst Rev 2014;5:CD003968. [doi: 10.1002/14651858.CD003968.pub4] [Medline: 24796681]
13. Sheinfeld Gorin S, Krebs P, Badr H, Janke EA, Jim HS, Spring B, et al. Meta-analysis of psychosocial interventions to reduce pain in patients with cancer. J Clin Oncol 2012 Feb 10;30(5):539-547 [FREE Full text] [doi: 10.1200/JCO.2011.37.0437] [Medline: 22253460]
14. O'Neil CK, Hanlon JT, Marcum ZA. Adverse effects of analgesics commonly used by older adults with osteoarthritis: focus on non-opioid and opioid analgesics. Am J Geriatr Pharmacother 2012 Dec;10(6):331-342 [FREE Full text] [doi: 10.1016/j.amjopharm.2012.09.004] [Medline: 23036838]
15. Sarzi-Puttini P, Cimmino MA, Scarpa R, Caporali R, Parazzini F, Zaninelli A, et al. Osteoarthritis: an overview of the disease and its treatment strategies. Semin Arthritis Rheum 2005 Aug;35(1 Suppl 1):1-10. [doi: 10.1016/j.semarthrit.2005.01.013] [Medline: 16084227]
16. Gunter RW, Whittal ML. Dissemination of cognitive-behavioral treatments for anxiety disorders: Overcoming barriers and improving patient access. Clin Psychol Rev 2010 Mar;30(2):194-202. [doi: 10.1016/j.cpr.2009.11.001] [Medline: 19942331]
17. Shafran R, Clark DM, Fairburn CG, Arntz A, Barlow DH, Ehlers A, et al. Mind the gap: Improving the dissemination of CBT. Behav Res Ther 2009 Nov;47(11):902-909. [doi: 10.1016/j.brat.2009.07.003] [Medline: 19664756]
18. Bennett GG, Glasgow RE. The delivery of public health interventions via the Internet: actualizing their potential. Annu Rev Public Health 2009;30:273-292. [doi: 10.1146/annurev.publhealth.031308.100235] [Medline: 19296777]
19. Rini C, Williams DA, Broderick JE, Keefe FJ. Meeting them where they are: Using the Internet to deliver behavioral medicine interventions for pain. Transl Behav Med 2012 Mar;2(1):82-92 [FREE Full text] [doi: 10.1007/s13142-011-0107-2] [Medline: 22924084]
20. 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]
21. 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]
22. Cuijpers P, Straten A, Andersson G. Internet-administered cognitive behavior therapy for health problems: a systematic review. Journal of Behavioral Medicine 2007;31(2):169-177. [doi: 10.1007/s10865-007-9144-1]
23. Cavanagh K, Shapiro DA. Computer treatment for common mental health problems. J Clin Psychol 2004 Mar;60(3):239-251. [doi: 10.1002/jclp.10261] [Medline: 14981789]
24. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]

XSL•FO
RenderX

25. Macea DD, Gajos K, Daglia Calil YA, 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]
26. Williams DA, Kuper D, Segar M, Mohan N, Sheth M, Clauw DJ. Internet-enhanced management of fibromyalgia: a randomized controlled trial. Pain 2010 Dec;151(3):694-702 [FREE Full text] [doi: 10.1016/j.pain.2010.08.034] [Medline: 20855168]
27. Pal K, Eastwood SV, Michie S, Farmer A, Barnard M, Peacock R, et al. Computer-based diabetes self-management interventions for adults with type 2 diabetes mellitus. Cochrane Database Syst Rev 2013;3:CD008776. [doi: 10.1002/14651858.CD008776.pub2] [Medline: 23543567]
28. Beatty L, Lambert S. A systematic review of internet-based self-help therapeutic interventions to improve distress and disease-control among adults with chronic health conditions. Clin Psychol Rev 2013 Jun;33(4):609-622. [doi: 10.1016/j.cpr.2013.03.004] [Medline: 23603521]
29. Civljak M, Stead LF, Hartmann-Boyce J, Sheikh A, Car J. Internet-based interventions for smoking cessation. Cochrane Database Syst Rev 2013;7:CD007078. [doi: 10.1002/14651858.CD007078.pub4] [Medline: 23839868]
30. Gallagher MW, Payne LA, White KS, Shear KM, Woods SW, Gorman JM, et al. Mechanisms of change in cognitive behavioral therapy for panic disorder: the unique effects of self-efficacy and anxiety sensitivity. Behav Res Ther 2013 Nov;51(11):767-777. [doi: 10.1016/j.brat.2013.09.001] [Medline: 24095901]
31. Goldin PR, Ziv M, Jazaieri H, Werner K, Kraemer H, Heimberg RG, et al. Cognitive reappraisal self-efficacy mediates the effects of individual cognitive-behavioral therapy for social anxiety disorder. J Consult Clin Psychol 2012 Dec;80(6):1034-1040 [FREE Full text] [doi: 10.1037/a0028555] [Medline: 22582765]
32. Turner JA, Holtzman S, Mancl L. Mediators, moderators, and predictors of therapeutic change in cognitive-behavioral therapy for chronic pain. Pain 2007 Feb;127(3):276-286. [doi: 10.1016/j.pain.2006.09.005] [Medline: 17071000]
33. Tasman A, Maj M, First M, Kay J, Lieberman J. Cognitive and Behavioral Therapies. In: Psychiatry. Chichester, West Sussex: Wiley-Blackwell; 2008.
34. Turk DC. Cognitive-behavioral approach to the treatment of chronic pain patients. Reg Anesth Pain Med 2003;28(6):573-579. [Medline: 14634950]
35. Keefe FJ, Caldwell DS, Williams DA, Gil K, Mitchell D, Robertson C, et al. Pain coping skills training in the management of osteoarthritic knee pain: A comparative study. Behavior Therapy 1990 Dec;21(1):49-62. [doi: 10.1016/S0005-7894(05)80188-1]
36. Keefe FJ, Blumenthal J, Baucom D, Affleck G, Waugh R, Caldwell DS, et al. Effects of spouse-assisted coping skills training and exercise training in patients with osteoarthritic knee pain: a randomized controlled study. Pain 2004 Aug;110(3):539-549. [doi: 10.1016/j.pain.2004.03.022] [Medline: 15288394]
37. Keefe FJ, Kashikar-Zuck S, Opiteck J, Hage E, Dalrymple L, Blumenthal JA. Pain in arthritis and musculoskeletal disorders: the role of coping skills training and exercise interventions. J Orthop Sports Phys Ther 1996 Oct;24(4):279-290. [doi: 10.2519/jospt.1996.24.4.279] [Medline: 8892142]
38. Porter LS, Keefe FJ, Garst J, Baucom DH, McBride CM, McKee DC, et al. Caregiver-assisted coping skills training for lung cancer: results of a randomized clinical trial. J Pain Symptom Manage 2011 Jan;41(1):1-13 [FREE Full text] [doi: 10.1016/j.jpainsymman.2010.04.014] [Medline: 20832982]
39. Keefe FJ, Shelby RA, Somers TJ, Varia I, Blazing M, Waters SJ, et al. Effects of coping skills training and sertraline in patients with non-cardiac chest pain: a randomized controlled study. Pain 2011 Apr;152(4):730-741 [FREE Full text] [doi: 10.1016/j.pain.2010.08.040] [Medline: 21324590]
40. Gil KM, Carson JW, Sedway JA, Porter LS, Schaeffer JJ, Orringer E. Follow-up of coping skills training in adults with sickle cell disease: analysis of daily pain and coping practice diaries. Health Psychol 2000 Jan;19(1):85-90. [Medline: 10711591]
41. Zautra AJ, Davis MC, Reich JW, Nicassario P, Tennen H, Finan P, et al. Comparison of cognitive behavioral and mindfulness meditation interventions on adaptation to rheumatoid arthritis for patients with and without history of recurrent depression. J Consult Clin Psychol 2008 Jun;76(3):408-421. [doi: 10.1037/0022-006X.76.3.408] [Medline: 18540734]
42. Sullivan MJ, Thorn B, Haythornthwaite JA, Keefe F, Martin M, Bradley LA, et al. Theoretical perspectives on the relation between catastrophizing and pain. Clin J Pain 2001 Mar;17(1):52-64. [Medline: 11289089]
43. Nielson WR, Jensen MP, Karsdorp PA, Vlaeyen JW. Activity pacing in chronic pain: concepts, evidence, and future directions. Clin J Pain 2013 May;29(5):461-468. [doi: 10.1097/AJP.0b013e3182608561] [Medline: 23247005]
44. Keefe FJ, Beaupre PM, Gil KM, Rumble ME, Aspnes AK. Group therapy for patients with chronic pain. In: Turk DC, Gatchel RJ, editors. Pain management: A practitioner's handbook. New York: Guilford Press; 2002:234-255.
45. Melzack R, Wall PD. Pain mechanisms: a new theory. Science 1965 Nov 19;150(3699):971-979. [Medline: 5320816]
46. Bandura A. Social foundations of thought and action: a social cognitive theory. Englewood Cliffs, NJ: Prentice-Hall; 1986.
47. Strecher V. Internet methods for delivering behavioral and health-related interventions (eHealth). Annu Rev Clin Psychol 2007;3:53-76. [doi: 10.1146/annurev.clinpsy.3.022806.091428] [Medline: 17716048]
48. U.S. Department of Health and Human Services. The research-based web design & usability guidelines. Washington, DC: U.S. Government Printing Office; 2006.

XSL•FO
RenderX

49. Bandura A. Self-efficacy: The exercise of control. New York: W. H. Freeman and Company; 1997.
50. Knowles MS. The adult learner: A neglected species. 3rd ed. Houston, TX: Gulf Pub. Co., Book Division; 1984.
51. Knowles MS, Holton EF, Swanson RA. The adult learner: The definitive classic in adult education and human resource development. 7th ed. Boston, MA: Elsevier; 2011.
52. Mayer RE. Applying the science of learning: evidence-based principles for the design of multimedia instruction. Am Psychol 2008 Nov;63(8):760-769. [doi: 10.1037/0003-066X.63.8.760] [Medline: 19014238]
53. Keefe FJ, Somers TJ. Psychological approaches to understanding and treating arthritis pain. Nat Rev Rheumatol 2010 Apr;6(4):210-216. [doi: 10.1038/nrrheum.2010.22] [Medline: 20357790]
54. Addis ME, Carpenter KM. The treatment rationale in cognitive behavioral therapy: Psychological mechanisms and clinical guidelines. Cognitive and Behavioral Practice 2000;7(2):147-156.
55. Addis ME, Jacobson NS. A closer look at the treatment rationale and homework compliance in cognitive-behavioral therapy for depression. Cognitive Therapy and Research 2000;24(3):313-326.
56. Bickmore T, Picard R. Establishing and Maintaining Long-Term Human-Computer Relationships. ACM Transactions on Computer Human Interaction 2005;12(2):293-327.
57. Bickmore T, Schulman D, Yin L. Maintaining engagement in long-term interventions with relational agents. Appl Artif Intell 2010 Jul 1;24(6):648-666 [FREE Full text] [doi: 10.1080/08839514.2010.492259] [Medline: 21318052]
58. Mayer RE. Principles of multimedia learning based on social cues: Personalization, voice, and image principles. In: The Cambridge handbook of multimedia learning. New York, NY: Cambridge University Press; 2005:201-212.
59. Mayer RE, DaPra CS. An embodiment effect in computer-based learning with animated pedagogical agents. J Exp Psychol Appl 2012 Sep;18(3):239-252. [doi: 10.1037/a0028616] [Medline: 22642688]
60. Reeves B, Nass C. The media equation: How people treat computers, television, and new media like real people and places. New York, NY: Cambridge University Press; 1996.
61. Bender JL, Radhakrishnan A, Diorio C, Englesakis M, Jadad AR. Can pain be managed through the Internet? A systematic review of randomized controlled trials. Pain 2011 Aug;152(8):1740-1750. [doi: 10.1016/j.pain.2011.02.012] [Medline: 21565446]
62. Schweiger Gallo I, Gollwitzer PM. Implementation intentions: a look back at fifteen years of progress. Psicothema 2007 Feb;19(1):37-42. [Medline: 17295981]
63. Andersson G, Paxling B, Wiwe M, Vernmark K, Felix CB, Lundborg L, et al. Therapeutic alliance in guided internet-delivered cognitive behavioural treatment of depression, generalized anxiety disorder and social anxiety disorder. Behav Res Ther 2012 Sep;50(9):544-550. [doi: 10.1016/j.brat.2012.05.003] [Medline: 22728647]
64. Knaevelsrud C, Maercker A. Does the quality of the working alliance predict treatment outcome in online psychotherapy for traumatized patients? J Med Internet Res 2006;8(4):e31 [FREE Full text] [doi: 10.2196/jmir.8.4.e31] [Medline: 17213049]
65. Preschl B, Maercker A, Wagner B. The working alliance in a randomized controlled trial comparing online with face-to-face cognitive-behavioral therapy for depression. BMC Psychiatry 2011;11:189 [FREE Full text] [doi: 10.1186/1471-244X-11-189] [Medline: 22145768]
66. Ritterband LM, Thorndike FP, Cox DJ, Kovatchev BP, Gonder-Frederick LA. A behavior change model for internet interventions. Ann Behav Med 2009 Aug;38(1):18-27 [FREE Full text] [doi: 10.1007/s12160-009-9133-4] [Medline: 19802647]
67. Hilgart M, Ritterband L, Thorndike F, Kinzie M. Using instructional design process to improve design and development of Internet interventions. J Med Internet Res 2012;14(3):e89 [FREE Full text] [doi: 10.2196/jmir.1890] [Medline: 22743534]
68. 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]
69. Rosser BA, McCullagh P, Davies R, Mountain GA, McCracken L, Eccleston C. Technology-mediated therapy for chronic pain management: the challenges of adapting behavior change interventions for delivery with pervasive communication technology. Telemed J E Health 2011 Apr;17(3):211-216. [doi: 10.1089/tmj.2010.0136] [Medline: 21375412]
70. Mayer RE. Rote versus meaningful learning. Theory Into Practice 2002 Nov;41(4):226-232.

Abbreviations

CBT: cognitive behavioral therapy
OA: osteoarthritis
PCST: Pain Coping Skills Training

XSL•FO
RenderX

Edited by G Eysenbach; submitted 03.03.14; peer-reviewed by M Buhrman, A Heapy, D Higgins2; comments to author 17.07.14; revised version received 29.08.14; accepted 02.10.14; published 19.12.14.

Please cite as:
Rini C, Porter LS, Somers TJ, McKee DC, Keefe FJ
Retaining Critical Therapeutic Elements of Behavioral Interventions Translated For Delivery via the Internet: Recommendations and an Example Using Pain Coping Skills Training
J Med Internet Res 2014;16(12):e245
URL: http://www.jmir.org/2014/12/e245/
doi:10.2196/jmir.3374
PMID:25532216

©Christine Rini, Laura S Porter, Tamara J Somers, Daphne C McKee, Francis J Keefe. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 19.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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

How Doctors View and Use Social Media: A National Survey

James Brown[1], BA, MPH, MBBS; Christopher Ryan[1], MBBS; Anthony Harris[1], MBBS, PhD

Western Clinical School, Department of Psychiatry, University of Sydney, Westmead, Australia

Corresponding Author:
James Brown, BA, MPH, MBBS
Western Clinical School
Department of Psychiatry
University of Sydney
Darcy Road
Westmead, 2145
Australia
Phone: 61 438798207
Fax: 61 298455000
Email: j.h.w.brown@gmail.com

Abstract

Background: Doctors are uncertain of their ethical and legal obligations when communicating with patients online. Professional guidelines for patient-doctor interaction online have been written with limited quantitative data about doctors' current usage and attitudes toward the medium. Further research into these trends will help to inform more focused policy and guidelines for doctors communicating with patients online.

Objective: The intent of the study was to provide the first national profile of Australian doctors' attitudes toward and use of online social media.

Methods: The study involved a quantitative, cross-sectional online survey of Australian doctors using a random sample from a large representative database.

Results: Of the 1500 doctors approached, 187 participated (12.47%). Most participants used social media privately, with only one-quarter not using any social media websites at all (48/187, 25.7%). One in five participants (30/155, 19.4%) had received a "friend request" from a patient. There was limited use of online communication in clinical practice: only 30.5% (57/187) had communicated with a patient through email and fewer than half (89/185, 48.1%) could offer their patients electronic forms of information if that were the patients' preference. Three in five participants (110/181, 60.8%) reported not being uncomfortable about interacting with patients who had accessed personal information about them online, prior to the consultation. Most of the participants (119/181, 65.8%) were hesitant to immerse themselves more fully in social media and online communication due to worries about public access and legal concerns.

Conclusions: Doctors have different practices and views regarding whether or how to communicate appropriately with patients on the Internet, despite online and social media becoming an increasingly common feature of clinical practice. Additional training would assist doctors in protecting their personal information online, integrating online communication in patient care, and guidance on the best approach in ethically difficult online situations.

(J Med Internet Res 2014;16(12):e267) doi:10.2196/jmir.3589

KEYWORDS

social media; Internet; professional practice; health communication; ethics; health policy; patient-physician relations

Introduction

There is widespread international disquiet about the ramifications of online and social media on clinical practice [1-7]. Doctors are uncertain of patient expectations [8,9], and of their ethical and legal obligations when using online communication [10,11]. However, there has been limited quantitative research into doctors' usage of this technology or their attitudes toward it [12-14]. As a consequence, professional guidelines aimed at providing ethical and practical oversight have been written with little empirically derived data.

The current standards of professional communication in medicine were developed with traditional face-to-face consultations as the template. This model of professionalism

XSL•FO
RenderX

has not transferred easily to an online environment and doctors are being left to act using their own intuition as new online ethical dilemmas arise [15-17]. For example, there is uncertainty as to how a doctor ought to respond to a "friend request" from a patient via social media, despite poor handling of the situation potentially affecting the viability of the therapeutic relationship [18]. Intuitions may differ between different age cohorts. Younger doctors, who use social media frequently [19], may be more comfortable communicating online [20,21] than senior doctors, who have limited familiarity with the technology.

While concerns about online searching [22,23], privacy [24], and professionalism [25] have been suggested as barriers to the use of social media by doctors, the slow adoption of social media also represents a possible lost opportunity for the medical profession at large to engage with patients and the general community. There is a need to examine the impact of the use of social media on physicians' knowledge, attitudes, skills, and behaviors in practice [26] and the influence of concerns about ethics, professionalism, and privacy upon that use [25,27].

The aim of our study was to investigate and quantify the current use of social and online media by doctors and gauge attitudes toward possible professional dilemmas when communicating with patients online in order to assist in the improvement of social media guidelines for effective use in clinical practice.

Methods

Participants

We selected a random sample of 1500 medical practitioners, who had email addresses on the Australasian Medical Publishing Company (AMPCo) database. The AMPCo database holds the contact details of 65,536 doctors, representing 93% of all doctors in Australia [28]. A total of 49% of the database had email addresses. AMPCo extracted all names using an automated randomization process that selects contacts proportionally from their database by age, gender, and location. Study investigators were unaware of the identity of any participant. This sample size was selected following review of expected response rates to unsolicited, non-incentivized, online medical research surveys [27,29-32] and in consultation with other large Australian-based surveys.

All participants were sent an invitation email including a link to an online survey and a participant information sheet explaining the rationale for the study. A repeat email was sent to non-responders 4 weeks later. No financial or educational incentives were provided to participants. Data was collected for 8 weeks from October to December 2013. All responses were automatically recorded via the survey platform and downloaded into an SPSS database. The study was approved by the University of Sydney Human Research Ethics Committee.

Questionnaire

The survey questionnaire was developed after a literature review following the Checklist for Reporting Results of Internet E-Surveys (CHERRIES) guideline [33]. It covered five broad areas of patient-doctor online interaction: (1) current participant usage, (2) general online behavior, (3) doctors' personal information online, (4) patients' information online, and (5) appropriate patient-doctor online interaction. Questions were multiple choice. A draft questionnaire was piloted and revised to a final survey of 36 items (plus demographic information) that took approximately 10 minutes to complete (see Multimedia Appendix 1).

Analysis

Data were analyzed with IBM SPSS Statistics 20 statistical software. Statistical approach was based upon distribution of responses. We compared proportions using chi-square analysis; means were compared using *t* tests. Logistic regression was used to estimate the association between social media use and doctor profile.

Results

Participants

Of the 1500 doctors who received the invitation email, 190 logged on to the survey and 187 completed it yielding a response rate of 12.47%. The participants were drawn from all states and territories of Australia and were representative of a broad range of specialties. The demographics of the participants were largely similar to that of the AMPCo database (Table 1), although a larger than expected proportion of participants (33.1% vs 23% [34]) identified as working in a rural location. The survey responses were generally consistent across all demographic groups. We have noted the exceptions to this in the text below.

XSL•FO
RenderX

Table 1. Participant demographics (n=187).

Demographics		n (%)
Current age (in years)		
	Less than 25	2 (1.1)
	25-35	43 (23.0)
	36-45	40 (21.4)
	46-55	45 (24.1)
	56-65	29 (15.5)
	66-75	11 (5.9)
	Greater than 75	5 (2.7)
	Missing	12 (6.4)
Gender		
	Male	95 (50.8)
	Female	80 (42.8)
	Missing	12 (6.4)
Rural or metropolitan		
	Rural	58 (31.0)
	Metropolitan	117 (62.6)
	Missing	12 (6.4)
Professional role		
	Intern / junior medical officer	17 (9.0)
	Doctors undertaking specialty training	36 (19.2)
	Specialist physicians	72 (38.5)
	General practitioner	46 (24.6)
	Not presently practicing	1 (0.6)
	Retired	2 (1.1)
	Other	1 (0.6)
	Missing	12 (6.4)
Private or public practice		
	Private only	65 (34.8)
	Public only	71 (38.0)
	Both public and private	37 (19.8)
	Not applicable	2 (1.1)
	Missing	12 (6.4)
Years since graduation		
	Less than 10	48 (25.7)
	10-19	39 (20.9)
	20-29	37 (19.8)
	30-39	37 (19.8)
	40-49	9 (4.8)
	Greater than 50	5 (2.7)
	Missing	12 (6.4)

XSL•FO
RenderX

Current Online and Social Media Use by Doctors

Most participants used social media at least once a week and only one-quarter did not use any social media websites at all (48/187, 25.7%). The most commonly used platform was Facebook (112/187, 59.9%). Most participants (147/187, 78.6 %) used social media in non-work hours, and of those, 16.3% (24/147) used social media more than 1 hour per day. In comparison, 38.0% (71/187) of participants used social media in work hours, with only 4.2% of those using the technology more than 1 hour per day.

There was a linear relationship between increasing age and decreasing social media use (OR 10.3, 95% CI 2.8-42.4). All interns, junior medical officers, and doctors undertaking specialty training used some form of social media compared with 72.2% (52/72) of specialist physicians and 69.6% (32/46) of general practitioners (*P*<.001).

Attitudes Toward, and Utilization of, Online Patient-Doctor Communication

Attitudes toward the use of social media with patients were divided. Although 67.0% (124/185) of participants agreed it might be appropriate for a doctor to interact with their patient via email, only 30.5% (57/187) of participants volunteered that they had done so. Only 1 of the 187 respondents had used social media (eg, Twitter or Facebook) to communicate with patients and only 21.2% (38/179) believed it would be appropriate to do so. Over one-third (63/185, 34.1%) of participants did not have a website or online presence for their practice, and over half (96/185, 51.9%) could not offer their patients electronic forms of information if that was a patient's preference.

In contrast to these low rates of online communication, doctors frequently spoke to their patients about online resources; 69.7% (129/185) had discussed online information sources (such as websites about their disease), though fewer had discussed social media resources such as online support groups (73/185, 39.5%). General practitioners were far more likely than specialist physicians to speak with their patients about Internet usage and online resources (44/46, 95.7% vs 43/72, 59.7%, *P*=.005). Rural and regional doctors were also more likely to have discussed Internet resources with their patients than their urban colleagues (45/58, 77.6% vs 77/117, 65.8%, *P*=.006).

Patient-Doctor Interaction on Facebook

Despite the very low rates of social media use in a professional setting, there were high rates of use in private with women being much more likely to use Facebook than men (59/80, 73.8% vs 49/95, 51.6%, *P*=.03). One in five participants (30/155, 19.4%) had received a friend request from a patient they only knew and interacted with professionally. When asked how they would respond to a patient who had sent them a friend request, participants were split in their responses (Table 2). Most commonly they would decline the request and do nothing more (54/155, 34.8%). Only 2.6% (4/155) would accept the friend request. Over half (89/155, 57.4%) thought it appropriate for a doctor to maintain a personal Facebook profile, though less than one-quarter (37/155, 23.9%) were comfortable with a patient being able to access content about the doctor posted on that page (such as photos posted by others).

Table 2. Doctors' response to friend request from a patient (n=155).

How would you respond to a patient who sent you a friend request on Facebook?	n (%)
Accept the request	4 (2.6)
Decline the request and do nothing more	54 (34.8)
Decline the request and send a private message explaining why	18 (11.6)
Decline the request and discuss at the next consultation	44 (28.4)
Do nothing	35 (22.6)

Protection of Personal Information Online

Most participants (110/181, 60.8%) reported they would not be comfortable interacting with a patient who had accessed personal information about them online prior to the consultation and 17.1% (31/181) of participants had experienced someone else posting information online about them, which they would not want patients to see. Although not common, some participants had interacted with patients who described information about them which they had not made available and which the patient found online (15/181, 8.3%) or on social media (3/181, 1.7%).

Few doctors were able to adequately protect their information online. While most participants were aware of the results that appear when they searched on the Internet for their full name (117/181, 64.6%), and the majority (107/155, 69.0%) have adjusted privacy settings to limit access to their information, a much smaller proportion took measures to control their online profile (65/181, 35.9%). The older participants were, the less likely they were to know how to remove photos of themselves they wouldn't want patients to see. No participant aged over 65 years knew how to do this, compared with 7% (3/45) aged 46-55 years, and 50% (1/2) aged under 25 years (*P*=.01). Females were more likely than males to control and curate their online profile, including adjusting privacy settings (38/80, 48% vs 25/95, 26%, *P*=.05).

Most participants (119/181, 65.8%) were hesitant to immerse themselves more fully in social media and online communication due to worries about public access and legal concerns.

Other Ethical Dilemmas

Doctors were also unsure if they had a duty to rebut inappropriate or inaccurate information posted online; with 38.1% (69/181) saying that doctors did and 29.3% (53/181) saying they did not, with the rest being undecided. There was no consensus about the appropriateness of accessing publically

XSL•FO
RenderX

available information about a patient and whether to broach that with the patient (Table 3).

Even in an emergency, 26.7% (48/180) of participants would not use publicly available online information (eg, a patient's Facebook page for information regarding a suicide attempt). Despite the split responses, 16.1% (29/180) of participants had already searched for information about a patient online (Table 4).

Table 3. Doctors' use of publically available patient information (n=180).

If you were to use publicly available online information about a patient to assist in their treatment, would you, as the doctor, discuss it with the patient?	n (%)
Yes, always	72 (40.0)
Yes, sometimes	18 (8.9)
Yes, rarely	4 (2.2)
No	4 (2.2)
Unsure	37 (20.6)
I would not use publicly available online information	47 (26.1)

Table 4. Examples of online ethical dilemmas.

Dilemma	Yes	No	Unsure
	n (%)	n (%)	n (%)
Do doctors have a duty to rebut inappropriate or inaccurate health information posted online? eg, a blogger saying that sex without a condom is safe.	69 (38.1)	53 (29.3)	59 (32.6)
Have you at any time searched for publicly available online information about a patient? eg, "Googled" a patient to find more information about them.	29 (16.1)	150 (83.3)	1 (0.6)
Is it appropriate for doctors to look up publicly available online information about a patient in an emergency? eg, searching a patient's Facebook page for information following a suicide attempt.	68 (37.8)	48 (26.7)	64 (35.6)
Is it appropriate for doctors to look up publicly available online information about a patient as part of regular clinical practice? eg, monitoring a pro-anorexia forum for posts made by one of your patients.	30 (16.7)	77 (42.8)	73 (40.6)

Discussion

Current Usage of Online and Social Media by Doctors is Limited

Our results confirm our hypothesis that Australian doctors have yet to fully integrate online communication and social media into their clinical practice, and many are unable to protect their personal information online.

Although Australian doctors frequently used social media in their own private lives, their use during their working day is minimal. This reflects similar usage patterns noted among US doctors [27]. Few changes have been made to integrate online communication into clinical practice (including resources as basic as email), despite this being a common expectation for client communication in most other professions. Over half of participants are unable to send information electronically; many practices cannot offer their patients the option of electronic communication at all, even if that is their preference. As a result, patients are not even given the choice of online communication in most medical facilities.

The relationship between age and social media use is not surprising. Younger doctors have grown up with online communication, and frequent personal use may have instilled confidence in their ability to navigate any potentially hazardous ethical dilemmas. In comparison, older doctors have not been as involved in the progressive integration of social media into daily life, nor the increasing volume of its use. As a result, social media fluency can vary greatly within a cohort of doctors within the same practice or hospital, interacting with the same patients.

In contrast to this low professional usage, many doctors are discussing Internet and social media resources with their patients. General practitioners (44/46, 96%) and rural doctors (44/58, 76%) report an extremely high rate of discussion about online resources, perhaps reflecting the central role of online health resources as part of primary care, especially in geographically distant centers, and the potential for further integrating online and social media in other areas of medical practice.

Doctors Are Unsure as to How to Respond to Online Ethical Dilemmas

Although patient-doctor online communication is currently limited, doctors are still encountering online ethical dilemmas. One in five doctors had already found themselves in situations where the traditional boundaries of the doctor-patient relationship had been stretched by a friend request on social media. This figure was considerably lower than the 34.5% of practicing physicians who reported receiving such a request from a patient reported in a recent US study [27]. When

XSL•FO
RenderX

questioned about how they would react in such a situation, Australian doctors were evenly distributed as to whether they would respond to or ignore the request and how they would do that, reflecting the absence of consensus on how best to approach such a situation.

Equally problematic was the question of whether it is appropriate to view publically available online information about patients, even in an emergency, and whether patients should be made aware that this information had been accessed. It is interesting to reflect upon whether a post on Facebook about a suicide attempt is any different to a written suicide note, as few would have any ethical concerns about reading the latter. Perhaps most surprisingly, given the readiness to discuss online resources with patients, doctors were unwilling to become involved in online discussions about the accuracy or appropriateness of online material or blogs. This may in turn be related to their uncertainty about the Internet and social media.

Doctors' Personal Information Is Not Adequately Protected

Doctors were concerned about legal issues when communicating with patients online and reported that privacy and legal concerns were driving their reluctance to participate more fully in social media. These concerns are comparable to those shown in other countries [27] and remain unaltered despite the development of guidelines for social media published by professional organizations before and during the study period [35-39]. The majority of doctors (125/181, 69.0%) stated they would be uncomfortable interacting with patients who had accessed online information about them prior to a consultation. Yet few take specific measures to manage and edit the information available about themselves online. Only one in five (39/181, 21.6%) know what to do if a compromising photo of them is posted without their permission, with 17.1% saying that information about them they wouldn't want patients to see had already been posted. This limited competency in managing the online information available about them means that the possible advantages of an Internet presence are lost and increases the risk of negative experiences, further reinforcing avoidance of the use of social media.

Many doctors believe they should not have a personal profile, to avoid such dilemmas coming about. This only further marginalizes doctors from developing experience and fluency in the use of the technology and ignores the fact that as time goes on a doctor's online profile will be developed even without their participation via the increasing use of third-party review sites [40,41] such as RateMDs.com and Google Reviews.

What is most evident from the results of this study is that the permeation of online and social media into everyday life is placing doctors in new situations that they find difficult to navigate. It is clear that the standards and practices that have previously guided everyday patient-doctor interactions are being placed in a new light as the profession adapts old understandings to new communication tools. For some, social media may simply be another innocuous form of communication, like a phone call or a text message. While others may consider the personal information about themselves available on social media to be only appropriate for their closest friends and family.

Limitations of the Research

The findings of the study are weakened by the modest response rate of 12.5%, though this is similar to that obtained by other studies using online collection [27,42]. As we only invited doctors via email this will have removed a number of participants who do not have an email address (51% of the AMPCo database do not have recorded email addresses). It is also likely that we elicited responses from doctors who are more interested in social media, potentially biasing our sample. This is likely to be in the direction of a group more cognizant with the use of the Internet and social media. Nevertheless, because our findings demonstrate a range of attitudes toward appropriate social media use with a population more engaged with social media, our study, if anything, likely underestimates doctors' discomfort with using social media.

Future Directions

This study reveals two key areas where research must be directed. First, consensus must be reached on how doctors should behave online. This requires an expanded investigation of global and country-specific standards for online communication and for professional bodies to make these expectations included as part of training and professional development. Second, current ethical and professional guidelines may be ineffective in guiding appropriate patient-doctor interaction online and informing doctors in how to protect their personal information. Therefore, a reappraisal of how these guidelines are promulgated needs to occur. Change in behavior in such a central part of medicine will require active engagement rather than the passive diffusion of guidelines. The researchers intend to repeat the study in 5 years and provide the survey tool to assist further social media research internationally.

Conclusion

Our study is a comprehensive description of current online and social media behavior of doctors. It underlines the continuing need to improve the online capabilities of doctors and refine online and social media guidelines for doctors, which have so far done little to improve the uncertainty of doctors online. It finds a surprising reluctance to engage with the new media despite the demands of the community. The results will also allow us to map emerging trends in social and electronic media use, bringing future ethical issues to light as online communication becomes more and more relevant to clinical practice.

XSL•FO
RenderX

Acknowledgments

Associate Professor Harris has received consultancy fees from Janssen Australia and Lundbeck Australia. He has received payments for educational sessions run for Janssen Australia and the Lundbeck Institute. He is or has recently been an investigator on industry-sponsored trials by Hoffman-La Roche, Janssen-Cilag Australia and Brain Resource Ltd.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Survey tool used in the research.

[PDF File (Adobe PDF File), 366KB - jmir_v16i12e267_app1.pdf]

References

1. Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. J Med Internet Res 2013;15(4):e85 [FREE Full text] [doi: 10.2196/jmir.1933] [Medline: 23615206]
2. Azu MC, Lilley EJ, Kolli AH. Social media, surgeons, and the Internet: an era or an error? Am Surg 2012 May;78(5):555-558. [Medline: 22546128]
3. Decamp M. Social media guidance, conflicts of interest, and health inequalities. Lancet 2012 Aug 4;380(9840):472-473. [doi: 10.1016/S0140-6736(12)61290-2] [Medline: 22863050]
4. Anonymous. Social media: how doctors can contribute. Lancet 2012 Apr 28;379(9826):1562. [doi: 10.1016/S0140-6736(12)60658-8] [Medline: 22541563]
5. Sweet M. Critics condemn proposals to restrict use of social media among health staff. BMJ 2012;345:e6240. [Medline: 22988309]
6. DeCamp M. Social media and medical professionalism: toward an expanded program. Arch Intern Med 2012 Oct 8;172(18):1418-1419. [doi: 10.1001/archinternmed.2012.3220] [Medline: 22911444]
7. Collier R. Professionalism: social media mishaps. CMAJ 2012 Sep 4;184(12):E627-E628 [FREE Full text] [doi: 10.1503/cmaj.109-4209] [Medline: 22802383]
8. Anikeeva O, Bywood P. Social media in primary health care: opportunities to enhance education, communication and collaboration among professionals in rural and remote locations: did you know? Practical practice pointers. Aust J Rural Health 2013 Apr;21(2):132-134. [doi: 10.1111/ajr.12020] [Medline: 23586577]
9. Antheunis ML, Tates K, Nieboer TE. 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]
10. St-Laurent-Gagnon T, Coughlin KW, Canadian Paediatric Society, Bioethics Committee. Paediatricians, social media and blogs: Ethical considerations. Paediatr Child Health 2012 May;17(5):267-272 [FREE Full text] [Medline: 23633902]
11. Payette MJ, Albreski D, Grant-Kels JM. "You'd know if you 'friended' me on Facebook": legal, moral, and ethical considerations of online social media. J Am Acad Dermatol 2013 Aug;69(2):305-307. [doi: 10.1016/j.jaad.2013.02.024] [Medline: 23866862]
12. 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;15(8):e184 [FREE Full text] [doi: 10.2196/jmir.2708] [Medline: 23985172]
13. Usher W. Types of social media (Web 2.0) used by Australian allied health professionals to deliver early twenty-first-century practice promotion and health care. Soc Work Health Care 2011;50(4):305-329. [doi: 10.1080/00981389.2010.534317] [Medline: 21512953]
14. Usher WT. Australian health professionals' social media (Web 2.0) adoption trends: early 21st century health care delivery and practice promotion. Aust J Prim Health 2012;18(1):31-41. [doi: 10.1071/PY10084] [Medline: 22394660]
15. Jain SH. Practicing medicine in the age of Facebook. N Engl J Med 2009 Aug 13;361(7):649-651. [doi: 10.1056/NEJMp0901277] [Medline: 19675328]
16. Gabbard GO, Kassaw KA, Perez-Garcia G. Professional boundaries in the era of the Internet. Acad Psychiatry 2011;35(3):168-174. [doi: 10.1176/appi.ap.35.3.168] [Medline: 21602438]
17. Mostaghimi A, Crotty BH. Professionalism in the digital age. Ann Intern Med 2011 Apr 19;154(8):560-562. [doi: 10.7326/0003-4819-154-8-201104190-00008] [Medline: 21502653]
18. Devi S. Facebook friend request from a patient? Lancet 2011 Apr 2;377(9772):1141-1142. [Medline: 21465700]
19. Moubarak G, Guiot A, Benhamou Y, Benhamou A, Hariri S. Facebook activity of residents and fellows and its impact on the doctor-patient relationship. J Med Ethics 2011 Feb;37(2):101-104. [doi: 10.1136/jme.2010.036293] [Medline: 21160080]
20. Chretien KC. Should I be "friends" with my patients on social networking web sites? Am Fam Physician 2011 Jul 1;84(1):105, 108 [FREE Full text] [Medline: 21766761]

XSL•FO
RenderX

21. Chretien KC, Farnan JM, Greysen SR, Kind T. To friend or not to friend? Social networking and faculty perceptions of online professionalism. Acad Med 2011 Dec;86(12):1545-1550. [doi: 10.1097/ACM.0b013e3182356128] [Medline: 22030752]
22. Kritz M, Gschwandtner M, Stefanov V, Hanbury A, Samwald M. Utilization and perceived problems of online medical resources and search tools among different groups of European physicians. J Med Internet Res 2013;15(6):e122 [FREE Full text] [doi: 10.2196/jmir.2436] [Medline: 23803299]
23. Samwald M, Kritz M, Gschwandtner M, Stefanov V, Hanbury A. Physicians searching the web for medical question answering: a European survey and local user studies. Stud Health Technol Inform 2013;192:1103. [Medline: 23920877]
24. Munson SA, Cavusoglu H, Frisch L, Fels S. Sociotechnical challenges and progress in using social media for health. J Med Internet Res 2013;15(10):e226 [FREE Full text] [doi: 10.2196/jmir.2792] [Medline: 24148206]
25. Grajales FJ, Sheps S, Ho K, Novak-Lauscher H, Eysenbach G. Social media: a review and tutorial of applications in medicine and health care. J Med Internet Res 2014;16(2):e13 [FREE Full text] [doi: 10.2196/jmir.2912] [Medline: 24518354]
26. McGowan BS, Wasko M, Vartabedian BS, Miller RS, Freiherr DD, Abdolrasulnia M. Understanding the factors that influence the adoption and meaningful use of social media by physicians to share medical information. J Med Internet Res 2012;14(5):e117 [FREE Full text] [doi: 10.2196/jmir.2138] [Medline: 23006336]
27. Bosslet GT, Torke AM, Hickman SE, Terry CL, Helft PR. The patient-doctor relationship and online social networks: results of a national survey. J Gen Intern Med 2011 Oct;26(10):1168-1174 [FREE Full text] [doi: 10.1007/s11606-011-1761-2] [Medline: 21706268]
28. Health Workforce Australia. Adelaide: HWA. Australia’s Health Workforce Series - Doctors in focus URL: https://www.hwa.gov.au/sites/uploads/australias_health_workforce_series_doctors_in_focus_20120322.pdf [accessed 2014-04-13] [WebCite Cache ID 6Onewgqha]
29. Sheehan KB. E-mail survey response rates: a review. Journal of Computer-Mediated Communication 2001;6(2). [doi: 10.1111/j.1083-6101.2001.tb00117.x]
30. Bonevski B, Magin P, Horton G, Foster M, Girgis A. Response rates in GP surveys - trialling two recruitment strategies. Aust Fam Physician 2011 Jun;40(6):427-430 [FREE Full text] [Medline: 21655493]
31. Shih T, Fan X. Comparing response rates from web and mail surveys: a meta-analysis. Field Methods 2008 Mar 03;20(3):249-271. [doi: 10.1177/1525822X08317085]
32. Kongsved S, Basnov M, Holm-Christensen K, Hjollund N. Response rate and completeness of questionnaires: a randomized study of Internet versus paper-and-pencil versions. J Med Internet Res 2007;9(3):e25 [FREE Full text] [doi: 10.2196/jmir.9.3.e25] [Medline: 17942387]
33. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
34. Australian Institute of Health and Welfare. Canberra: AIHW. 2008. Rural, regional and remote health: indicators of health system performance URL: http://www.aihw.gov.au/WorkArea/DownloadAsset.aspx?id=6442459852 [accessed 2014-04-13] [WebCite Cache ID 6Onetaphb]
35. Frankish K, Ryan C, Harris A. Psychiatry and online social media: potential, pitfalls and ethical guidelines for psychiatrists and trainees. Australas Psychiatry 2012 Jun;20(3):181-187. [doi: 10.1177/1039856212447881] [Medline: 22679216]
36. American Medical Association. 124 - Professionalism in the Use of Social Media. Opinion 9 URL: http://www.ama-assn.org//ama/pub/physician-resources/medical-ethics/code-medical-ethics/opinion9124.page [accessed 2014-06-08] [WebCite Cache ID 6QAW3fCjz]
37. Australian Medical Association. 2011. Social media and the medical profession: a guide to online professionalism for medical practitioners and medical students URL: https://ama.com.au/system/files/node/6231/Social%20Media%20and%20the%20Medical%20Profession_FINAL%20with%20links.pdf [accessed 2014-06-08] [WebCite Cache ID 6QAW8tfsY]
38. General Medical Council. Doctors' use of social media URL: http://www.gmc-uk.org/guidance/ethical_guidance/21186.asp [accessed 2014-08-17] [WebCite Cache ID 6Rt232PB1]
39. Farnan JM, Snyder Sulmasy L, 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]
40. 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;16(3):e97 [FREE Full text] [doi: 10.2196/jmir.3145] [Medline: 24686918]
41. Verhoef LM, Van de Belt TH, Engelen LJ, Schoonhoven L, Kool RB. Social media and rating sites as tools to understanding quality of care: a scoping review. J Med Internet Res 2014;16(2):e56 [FREE Full text] [doi: 10.2196/jmir.3024] [Medline: 24566844]

XSL•FO
RenderX

42. Scott A, Jeon SH, Joyce CM, Humphreys JS, Kalb G, Witt J, et al. A randomised trial and economic evaluation of the effect of response mode on response rate, response bias, and item non-response in a survey of doctors. BMC Med Res Methodol 2011;11:126 [FREE Full text] [doi: 10.1186/1471-2288-11-126] [Medline: 21888678]

Abbreviations

AMPCo: Australasian Medical Publishing Company

Edited by G Eysenbach; submitted 08.06.14; peer-reviewed by M Decamp, W Usher, W Glinkowski; comments to author 10.08.14; revised version received 01.09.14; accepted 20.10.14; published 02.12.14.

Please cite as:
Brown J, Ryan C, Harris A
How Doctors View and Use Social Media: A National Survey
J Med Internet Res 2014;16(12):e267
URL: http://www.jmir.org/2014/12/e267/
doi:10.2196/jmir.3589
PMID:25470407

©James Brown, Christopher Ryan, Anthony Harris. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 02.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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

Motivations for Contributing to Health-Related Articles on Wikipedia: An Interview Study

Nuša Farič[1*], BSc(Psych)(Hons), MSc; Henry WW Potts[1*], BA (Hons), MSc, PhD, Pg Cert

UCL, Centre for Health Informatics & Multiprofessional Education (CHIME), University College London, London, United Kingdom

[*]all authors contributed equally

Corresponding Author:
Nuša Farič, BSc(Psych)(Hons), MSc
UCL
Centre for Health Informatics & Multiprofessional Education (CHIME)
University College London
The Farr Institute of Health Informatics Research
222 Euston Road
London, NW1 2DA
United Kingdom
Phone: 44 (0)20 3549
Fax: 44 7679 8002
Email: nfaric@gmail.com

Abstract

Background: Wikipedia is one of the most accessed sources of health information online. The current English-language Wikipedia contains more than 28,000 articles pertaining to health.

Objective: The aim was to characterize individuals' motivations for contributing to health content on the English-language Wikipedia.

Methods: A set of health-related articles were randomly selected and recent contributors invited to complete an online questionnaire and follow-up interview (by Skype, by email, or face-to-face). Interviews were transcribed and analyzed using thematic analysis and a realist grounded theory approach.

Results: A total of 32 Wikipedians (31 men) completed the questionnaire and 17 were interviewed. Those completing the questionnaire had a mean age of 39 (range 12-59) years; 16 had a postgraduate qualification, 10 had or were currently studying for an undergraduate qualification, 3 had no more than secondary education, and 3 were still in secondary education. In all, 15 were currently working in a health-related field (primarily clinicians). The median period for which they have been an active editing Wikipedia was 3-5 years. Of this group, 12 were in the United States, 6 were in the United Kingdom, 4 were in Canada, and the remainder from another 8 countries. Two-thirds spoke more than 1 language and 90% (29/32) were also active contributors in domains other than health. Wikipedians in this study were identified as health professionals, professionals with specific health interests, students, and individuals with health problems. Based on the interviews, their motivations for editing health-related content were summarized in 5 strongly interrelated categories: education (learning about subjects by editing articles), help (wanting to improve and maintain Wikipedia), responsibility (responsibility, often a professional responsibility, to provide good quality health information to readers), fulfillment (editing Wikipedia as a fun, relaxing, engaging, and rewarding activity), and positive attitude to Wikipedia (belief in the value of Wikipedia). An additional factor, hostility (from other contributors), was identified that negatively affected Wikipedians' motivations.

Conclusions: Contributions to Wikipedia's health-related content in this study were made by both health specialists and laypeople of varying editorial skills. Their motivations for contributing stem from an inherent drive based on values, standards, and beliefs. It became apparent that the community who most actively monitor and edit health-related articles is very small. Although some contributors correspond to a model of "knowledge philanthropists," others were focused on maintaining articles (improving spelling and grammar, organization, and handling vandalism). There is a need for more people to be involved in Wikipedia's health-related content.

(J Med Internet Res 2014;16(12):e260) doi:10.2196/jmir.3569

XSL•FO
RenderX

KEYWORDS

motivation; writing; social media; wikis; Wikipedia; consumer health information

Introduction

Wikipedia is a free online encyclopedia created through the collaborative efforts of volunteers. As an open wiki, anyone can freely add to, modify, or delete its contents. It has become a major source of health-related information for health care professionals, students, patients, and the general public.

Wikipedia is the largest online encyclopedia with more than 26 million articles in 287 different languages and more than 4 million in the English-language Wikipedia [1]. It currently ranks as the sixth most-visited site on the Internet [2], attracting over 365 million unique visitors monthly or 29.5% of global Internet consumers [3,4]. Seeking health information online is now commonplace and widespread globally [5-11]. As of March 2013, the English-language Wikipedia contained more than 28,216 medical articles [12] and is a prominent repository of online health information. When health terms are searched in popular search engines such as Google and Yahoo, Wikipedia appears in the top 10 results 71%-85% of the time [13]. Wikipedia's global popularity as an online health resource has also been observed among physicians, with 70% reporting using it in 1 study [14].

Wikipedia has attracted controversy around the reliability of its entries [15]. A comparison of science-related topics with Encyclopaedia Britannica indicated a comparable error rate [16]. Czarnecka-Kujawa and colleagues [17] found Wikipedia's medical specialty entries were comprehensive (compared with ICD-9/10) and had moderate reliability. The risks associated with misinformation have raised the standards of control on Wikipedia where there are numerous policies, guidelines, and collaborative systems in place to ensure the quality of information [18]. The focus of these has been on biographical articles where there is a threat of libel litigation rather than on health-related articles, but there are many specific guidelines on medical topics. Although the reliability of Wikipedia is of great importance, it is not the primary field of enquiry for this study. We suggest that the debate around the reliability of information is too simple. Equally accurate articles can have different focuses, different styles of writing, and so on, all which may affect how useful articles are for different audiences and how different audiences use articles. The content of Wikipedia articles ultimately depends on those who contribute to Wikipedia (Wikipedians) and their reasons for doing so, and this is our focus.

There are 18 million registered accounts on Wikipedia. Several studies have examined the characteristics of Wikipedians: they are more familiar with the topics they edit than average Internet users [19] and more often male [20-22]. A 2011 Wikimedia survey of more than 5000 Wikipedians found an average age of 28 years, with 61% having a university degree, 18% a Master's, and 8% a Doctoral degree. Previous studies examined culturally bound differences in Wikipedia contribution [23] and found Wikipedians score significantly lower on agreeableness, conscientiousness, and openness in the Big Five personality traits compared to non-Wikipedians [24]. However, characteristics of those specifically editing health-related pages have not been described.

A number of studies have examined knowledge-sharing intention and behavior [25-27]. Knowledge-sharing behavior has traditionally been studied in an organizational context; recently its principles have been applied to electronic networks [28]. This study assumed that contribution to Wikipedia can be classified as knowledge-sharing behavior because individuals largely engage by contributing what they know to specific pages.

Research on knowledge-sharing behavior has revealed 2 classes of motivations. These are intrinsic motivations, such as internal feeling of enjoyment and satisfaction [29], and extrinsic or goal-directed motivations, such as obtaining a reward and reputation [30]. The expectation of either obtaining the internal gratification or extrinsic returns may motivate individuals. However, Wikipedia poses a departure from conventional modes of knowledge-sharing behavior. Wikipedia is the prime example of a "commons-based peer production" model [31,32]. Wikipedians do not receive monetary or formal awards for their voluntary contributions and their unconventional modes of engagement have received scholarly attention [33,34]. Therefore, conventional motivational theory may be circumscribed in the context of Wikipedia and it is individual personality differences that affect how the information is produced and used. For example, Kuznetsov [35] found that Wikipedians are motivated by a process of interrelated value systems, such as altruism, reciprocity, community, autonomy, and reputation. It would be reasonable to surmise that in Wikipedia the exchange of knowledge is not based on interpersonal relationships (ie, intrinsic, extrinsic gains), but on the relative merit and importance one assigns to the context in which knowledge is shared [20,30,35].

Recognizing the limitations of conventional motivational theory, Leonard et al [36] proposed a model of self-concept motivations: the individuals are motivated to perform a behavior based on their inherent standards (internal self-concept) or standards that are in accordance with a reference group (external self-concept). The individual's motivations stems either from meeting a set of perceptions of the self (ie, of their own values and competencies) or the perceptions of the ideal-self (ie, of values, competencies, and success of the reference group). Applying this to Wikipedia, Yang and Lai [20,30] observed that internal self-concept-based motivation is the chief motivation predictor in knowledge-sharing intention and behavior. Individuals were most likely to share knowledge in Wikipedia due to the confidence in their capabilities, affirmed by the concept of self-efficacy. Studies on organizational knowledge sharing have shown that self-efficacy is the crucial predictor of knowledge-sharing intention and behavior [25,37]. Yang and Lai [30] also revealed that the quality of the information and the quality of the information system yielded a positive attitude toward Wikipedia. Prior research of individual behavior has demonstrated that individual attitudes are good predictors of

XSL•FO
RenderX

behavior and have been found to be crucial in knowledge-sharing intention and behavior [21,26,27,35,38].

These prior studies suggest the motivations behind Wikipedia contributions, highlighting that motivated behavior rarely pertains to a single motivation [39]. Considering the lack of data on motivations of Wikipedia contributors, a growing number of health-related queries online and Wikipedia's status as a prominent resource for health information, it is important to answer the question of who contributes to the health-related Wikipedia pages and why.

Methods

Design

We employed a cross-sectional design. The study was conducted between May to September 2012. The recruited sample were Wikipedia users with an editing history in health-related entries in the English-language Wikipedia. The first part of the study used a questionnaire, whereas the second part used semi-structured interviews.

Ethics

The study was approved by the University College London (UCL) Ethics Committee and the Wikimedia Foundation Research Committee.

Article Sampling

To sample contributors of health-related articles, we first generated a sample of health-related articles on Wikipedia. It is difficult to sample from all health-related articles on Wikipedia. Wikipedia articles are sorted into categories, but this task is done in the same way as all Wikipedia editing and is incomplete. Articles vary in size, importance, and how often they are accessed. Defining what constitutes being health-related is also difficult, with many marginal cases.

An earlier study on Wikipedia extracted health-related keywords from 3 indexes of the online health service websites, namely MedlinePlus, National Health Service (NHS) Direct Online, and the National Organization of Rare Disorders (NORD), the last to oversample rarer conditions [12]. Their lists of keyword phrases consisted of 1726 items for MedlinePlus, 966 items for NHS Direct Online, and 1173 items for NORD. We randomly selected 11 keyword phrases from each of these lists. These were entered into the search box on Wikipedia and the best matching article chosen. In addition, the study used articles listed under the category of Selected Articles on the Wikipedia Portal Medicine [40], constituting a set of articles the Wikipedia community have chosen to highlight as being of high quality and interesting. The 11 most recently edited articles were chosen from the selected article list from between July 11-25, 2012. The final sampled articles list is shown in Table 1.

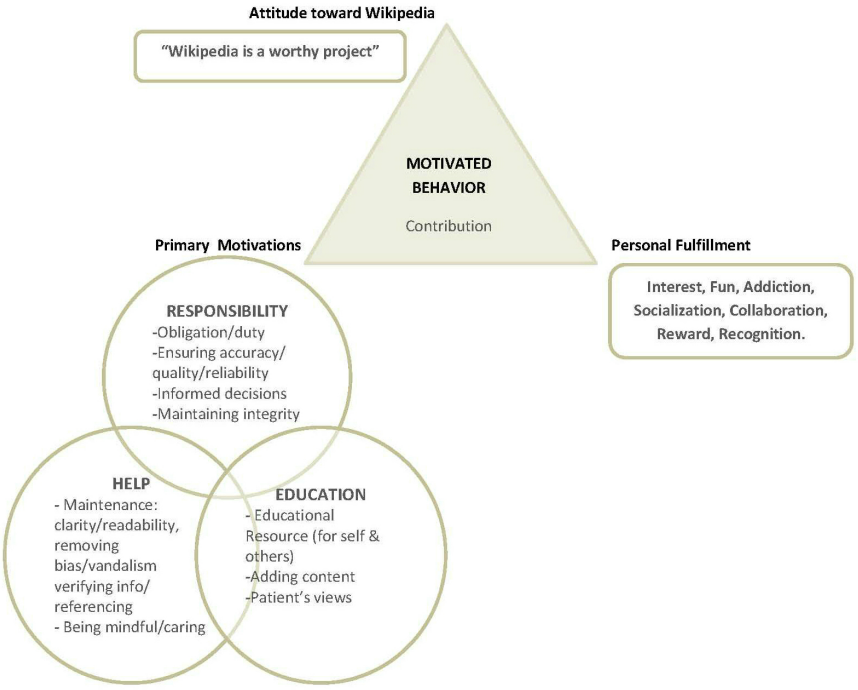
Having produced a sample of articles, a sample of contributors was produced by selecting the most recent 5 contributors for each article listed under the "history" tab.

Table 1. The sampled articles list for each of the medical databases and Wikipedia Portal Medicine. For the first 3 columns, the first term is the keyword phrase selection, whereas the bracketed term refers to the Wikipedia article name if it was not the same.

MedlinePlus	NHS Direct	NORD	Wikipedia Portal Medicine
Barrett's esophagus	Contact dermatitis	Very long chain acyl CoA dehydrogenase deficiency (very long-chain acyl-coenzyme A dehydrogenase deficiency)	Asthma
Menopausal hormone therapy (hormone replacement therapy (menopause))	Oral thrush (oral candidiasis)	Fiber type disproportion (congenital fiber type disproportion)	Insulin
Living wills (advance health-care directive)	Epidermolysis bullosa	Congenital fibrodysplasia ossificans progressive (fibrodsysplasia ossificans progressiva)	Helicobacter pylori
AMD (muscular degeneration)	Indigestion	Pancreatic islet cell tumor (pancreatic cancer)	Forensic facial reconstruction
Methamphetamine	Artificial insemination	Dubin Johnson Syndrome	Metabolism
Osteonecrosis (avascular necrosis)	Rectal examination	L1 syndrome (MASA syndrome)	Influenza
ERT (hormone replacement therapy)	Rheumatic fever	Leukodystrophy, metachromatic (metachromatic leukodystrophy)	Sexually transmitted disease
Staphylococcal infections (*Staphylococcus*)	Vitiligo	Cerebral palsy	Female hysteria
Arthrography	Nasal polyps (nasal polyp)	Irritable bowel syndrome	Vacutainer
Arm injuries and disorders (median nerve palsy)	Bulimia	Mesothelioma	Nutrition
Implantable defibrillators (implantable cardioverter defibrillator)	Psychotherapy	Cataracts (cataract)	2007 Bernard Mathews H5N1 outbreak

XSL•FO
RenderX

Figure 1. The motivational model.



Wikipedia users may register and create an account or they may edit articles without registering (or without logging into their registered account), in which case their edits are shown as coming from an Internet Protocol (IP) address. We included both registered (account) and nonregistered (IP address) contributors. Bots (automated or semiautomated software tools) were excluded. Every account and every editing IP address has or can have a Talk page, where a standardized invitation message was placed. NF set up a Wikipedia account for the purposes of recruitment [41].

A total of 44 articles were selected, which could have yielded up to 220 contributors to contact. In practice, many individuals came up more than once in the sampling frame, in which case, accounts not previously contacted were chosen until 220 different accounts were contacted.

The invitation message included brief information about the purpose of the study and the selected members were asked to follow a link that took them to NF's Wikimedia Commons page [41] containing instructions for participation and information about the study. The instructions asked participants to complete a Web-based questionnaire by clicking on the provided hyperlink. This was run through UCL's Opinio system [42].

Questionnaire

A 16-item questionnaire included questions referring to participants' characteristics, such as age, country of residence, employment, education, and Wikipedia editing history (including types of edits, editing in other languages, and number of health-related pages edited). The survey also included the question "What are your main motivations for editing health-related pages on Wikipedia?" with a free-text response. The penultimate item asked participants whether they would be willing to be interviewed and the final item asked for contact details (Wikipedia username or email address). A total of 32 complete survey responses were received (response rate 14.5%). Of these, 91% (29/32) agreed to be interviewed.

Participants

All participants were registered Wikipedians who had previously edited at least 1 health-related page on the English-language Wikipedia. A total of 32 Wikipedians (31 male and 1 female) volunteered to participate and completed the questionnaire. Due to time constraints, only 17 participants were interviewed through Skype, via email, or in person at UCL (in London). The reason for a much smaller proportion of participants interviewed than indicated is that they did not respond when contacted to be interviewed. All interviewees returned a signed and dated copy of the informed consent form. The age of the questionnaire respondents ranged from 12 to 59 years with a mean of 37 (SD 13) years. The mean age of the interview sample was 40 years.

Interviews

Overview

Two different semi-structured interview formats were used, but using the same interview schedule, developed around topics exploring participants' personal characteristics and individual motivations around editing Wikipedia. The questions were open-ended and were presented in the same order in both interview settings. The interview guides were developed by eliciting information around the topic of interest (ie, experience and motivations for editing) [43]. The interviews followed a general-to-specific approach and interviews were piloted before full use and slight amendments made following a reflexivity exercise [44].

Interview Schedule

The first interview schedule consisted of 25 questions that participants were asked to answer in a written format and return the completed answers to the researcher via email. The second schedule allowed participants to respond to the questions directly

XSL•FO
RenderX

either face-to-face or via Skype. Responses were recorded with a digital voice recorder. Both schedules incorporated questions (described subsequently). In most cases, the interview answers prompted issues requiring further exploration, which was done in an unstructured manner. If a participant responded by email, additional questions were sent via email. Face-to-face and Skype interviews lasted 30-120 minutes. Written responses contained from 969 to 3475 words. The list of questions used in the interview were:

1. What do you do and what are your specific interests?
2. What propelled you to start editing Wikipedia health-related pages?
3. Why did you edit the specific health-related page(s)?
4. Are your interests related to your Wikipedia edits?
5. Why do you edit with an account vs nonaccount?
6. What type of vandalism do you revert and why?
7. Do you have a particular group of readers in mind when you are editing health-related Wikipedia pages?
8. What are your main motivations for editing health-related content on Wikipedia?
9. Please talk about your reasons for editing health-related Wikipedia pages and comment on whether the reasons are entirely personal or driven by any external factors (such as a group of people, organization, or a particular individual)?
10. Are there any factors you can think of that would impact your motivations for editing health-related content on Wikipedia?
11. Some people say that one of their main motivations is to promote collaboration between patients, carers, and medical professionals. Is this the case for you?
12. What are your views on the quality of the health-related content on Wikipedia?

Data Analysis

A free-text section of the questionnaire was included and analyzed with the interview content. All interviews were transcribed verbatim. The transcripts were analyzed according to grounded theory realist analysis [44,45]. Grounded theory is a well-documented method to explore a concept within a specific context.

Transcripts were systematically coded line-by-line by NF until initial descriptive labels of motivations emerged. These labels were clustered to form a set of concepts that informed the coding paradigm. The new emerging codes were constantly checked retrospectively and prospectively against the higher order categories through constant comparative method [45,46]. NF's reflections were noted as memos, which also included instances of negative cases [44]. A sample (6/32) of interviews were independently coded by HP and 2 other health psychology researchers and no disagreements were noted.

Results

Summary

Almost half of the participants (47%, 15/32) reported currently working in a health-related field, generally as physicians. Of these, 6% (2/32) indicated that their edits were exclusively health-related. The rest of the health-related employment fields covered lung cancer research, health education, health psychology research, regulatory affairs, medical literature, chemistry research, pharmaceutical industry, and health-related advertising. Participants who indicated they were not currently employed in a health-related field (53%, 17/32) did not usually list their professions, but those who were interviewed worked in fields such as engineering, theology, literature, and a couple of participants were students (Table 2).

The reported edits for health-related pages ranged from 1 to more than 50,000. Edit counts do not necessarily reflect the types of edits performed. Of the sample, editorial activities were reported as 16% (5/32) performing primarily major edits, such as adding content and providing quality references, whereas 28% (9/32) reported performing primarily minor edits described as "maintenance issues" that included reverting vandalism, correcting errors, paragraphing, grammar and style, linking articles, checking sources, and simplifying prose. In all, 56% (18/32) reported performing both types of edits. The edit category was self-selected on a demographics questionnaire by respondents and further details provided in the free-comment section.

Country of residence included the United States (n=12), United Kingdom (n=6), Canada (n=4), Australia (n=2), and 1 each in Sweden, The Netherlands, France, Austria, Malaysia, South Africa, and Columbia.

A substantial proportion reported speaking more than 1 language (66%, 21/32), which included French (n=11), Spanish (n=6), German (n=4), Dutch (n=2), Swedish (n=2), and 1 participant each for Chinese, Italian, Afrikaans, Malay, and Bengali. Of the multilingual sample, 25% (8/32) also reported editing Wikipedia pages in these languages.

Approximately 90% (29/32) of Wikipedians were also active contributors in domains other than health, which included both very specific and general descriptors related to topics such as religion, languages, literature, history, sport, politics, architecture, engineering, pop culture, geology and mythology.

Emergent Categories

People contribute to Wikipedia in different capacities and for various different reasons. The method of realist grounded theory allowed for the emergence of an explanatory theoretical framework. Three interlinked themes that arose from the data were identified and labeled as "help," "education," and "responsibility." Each of these comprised a set of subcategories, shown in Figure 1. In addition, 2 further categories namely "personal fulfillment" and "attitude toward Wikipedia" had a significant motivating quality and were included in the resulting model.

The 3 core motivational systems result in a motivated behavior: contribution to Wikipedia. Contribution, which results in knowledge building or knowledge growth, is also Wikipedia's core mission and concept as "sum of all knowledge."

XSL•FO
RenderX

Table 2. Participant characteristics (N=32).

Variable	n (%)
Gender	
Male	31 (97)
Female	1 (3)
Age (band)	
10-20	4 (12)
21-30	6 (19)
31-40	11 (34)
41-50	4 (12)
51-60	7 (23)
Highest level of education	
Secondary school	3 (9)
Still at school	3 (9)
College/university degree (eg, BSc, BA)	8 (25)
Studying as an undergraduate student	2 (6)
Master's university degree (eg, MA, MSc)	6 (19)
Doctorate/professional degree (eg, PhD, MD)	10 (31)
Employment	
Yes-full time	20 (62)
Yes-part time	5 (16)
No	6 (19)
Retired	1 (3)
Currently working in a health-related field	
No	17 (53)
Yes	15 (47)
Wikipedia editing history	
<6 months	1 (3)
<1 year	3 (9)
1-2 years	5 (16)
3-5 years	8 (25)
5-8 years	10 (31)
≥8 years	5 (16)
Health-related pages edited	
<10	5 (16)
Approximately 10-20	5 (16)
Approximately 20-30	1 (3)
Approximately 30-50	2 (6)
Approximately 50-100	5 (16)
>100	14 (44)
Types of edits	
Mainly minor	9 (28)
Mainly major	5 (16)
Both	18 (56)

XSL•FO
RenderX

Motivation Help

All Wikipedians in the study shared a goal of using their skills in order to improve Wikipedia. This was expressed with phrases such as maintaining, ensuring, providing, building, removing, taking care of, adding, clearing, weeding, simplifying, verifying, sharing, expanding, fixing, and helping. These activities can be understood as encompassing both major and minor edits.

There was an almost unanimous answer to the question about what propelled the participants to begin editing Wikipedia, shown in this example:

> *I found a mistake and I discovered that I could fix it. That has propelled most of my work on the site since then.*

It was observed that the motivation help generally stemmed from a sense of importance and care that is also a characteristic of responsibility. In some cases, caring in turn implied being mindful, not just of the self in relation to Wikipedia, but also the relationships others may have with Wikipedia:

> *I've always liked the idea of being able to fix something and have many people be able to benefit from my efforts. I noticed that the...article needed attention, and since I knew about it, I thought I'd tackle the job of fixing it up. No one else seemed like they cared, but the article was read 500-600 times per day, and this bothered me.*

> *I will frequently work on 1 page and make hundreds of edits to it until it is brought up to the professional standard. That is where my prime activity is making major edits to significant disease-related articles. So, this last month I made about 500 edits on the article on...updating it to the most recent literature and I am working in collaboration with a group called translators without borders, to translate these key articles to as many other languages possible.*

Not all reported having health-related interests. Their editing behavior was expressed accordingly:

> *I edit mainly to remove vandalism...therefore, I edit whatever comes up in the queue, which may or may not be health related.*

> *Health-related pages in particular are frequently written from a perspective of a physician rather than a layman. They also require more reliable sources. I try to help solve both of these issues.*

Help was a pervasive motivation shown through a number of different editing behaviors. For a very small proportion of Wikipedians, motivation help was not implicit, but was expressed instead as activity reflective of their inherent traits of character:

> *I am a proofreader by vocation. I cannot leave bad grammar go uncorrected.*

Motivation Education

The motivation education was the most frequently emerging motivation. The initial decision to come to Wikipedia was reported as "intellectual curiosity," such as a need for information and a need for learning:

> *I just want to educate myself because I am interested and then share it with people around me.*

> *My edits were sustained by simple interest in topics; I had read a book or article on something and ended up on the Wikipedia page so I'd add a summary or comment about said topic.*

In the process of editing a Wikipedia article, the contributors are expected to follow and adhere to Wikipedia's guidelines, such as verifiability [47], neutral point of view (NPOV) [48], and the guidelines on provision of the evidence-based claims [49]. The process of actively looking for reliable sources has been described as educational:

> *...very quickly when I began editing Wikipedia I learned that I needed to adjust how I write because there is no argument from authority. It's all about the citations and sources that you can cite. So the verifiability of the Wikipedia was actually educational for me because I realized that sometimes I said things for which it was hard to find a reference.*

The egalitarian nature of Wikipedia, coupled with the guidelines on provision of the evidence-based claims, makes it expected for everyone to support their contributions with a reliable source, which is inherently an educational activity. The process of acquiring new knowledge leads to a better understanding of a specific topic which can be perceived as a form of personal gain (ie, personal fulfillment, which can also be applied across backgrounds):

> *I usually start by finding the most recent review articles in the medical literature that discuss this topic. Almost invariably, I learn things about conditions that I had previously been unaware of; I can think of a number of examples where my management of patients with a particular condition has been better because I had worked on the relevant Wikipedia article.*

As an online knowledge repository, Wikipedia is also a place for groups of people to exchange information. A number of Wikipedians reported that Wikipedia's international community added to their educational awareness:

> *Occasionally when I need to look up a rare disease, I'm almost afraid to look at the talk pages because one finds actual patients and their families pleading for help that isn't there yet...however, it's important, in fact, essential reading for the researchers working to find treatments for these conditions.*

The motivation education was ubiquitous for Wikipedians in the study and encompassed both teaching and learning, either in the process of reading Wikipedia, updating Wikipedia, or through intellectually challenging debates. The motivation education is cyclical and an inevitable state of the editing process on Wikipedia. In sum:

XSL•FO
RenderX

> *...the short answer to your question [of main motivation] is hedonistic intellectual enjoyment coupled with a sense of responsibility.*

Motivation Responsibility

The motivation responsibility was very closely tied to all other motivations, but the strength of responsibility depended on the attitude toward Wikipedia (ie, beliefs about Wikipedia and beliefs about health care). The goal of many of the contributors in this study was to communicate clearly presented and verifiable information to the world and that applied to the task for both major and minor contributors:

> *...as a physician, we take the Hippocratic Oath. We try to do the best we can for patients and I consider my patients to be all people globally. And to help all people globally, one way to do that is to provide them access to high quality health care information. So if I can't see them personally in my emergency department, I know that hundreds of millions of them are looking at Wikipedia to help answer their questions.*

> *I wanted to help the potential future readers who will consult those articles before or instead of a doctor. Their health care decisions may depend on the information they find. I'm no doctor, but at least I can make the articles easier to read.*

Several contributors recognized Wikipedia's scale of influence as the largest repository of online health information, accessible to the whole world, the successful delivery of health care is of vital importance:

> *Wikipedia is necessarily a distillation of many facts to the things that are most important and valuable. So participating in that is something I felt almost an obligation...I often have felt that everybody in the world should have access to the information and Wikipedia was one place that everyone could access and I couldn't think of another reference work that would be so useful.*

Thus, many Wikipedians reported taking on the responsibility of educating the public. Similarly, they felt responsible to educate their peers, family, friends, students, or colleagues, but were often met with resistance, perhaps something that further strengthened their sense of responsibility:

> *I want the material to be as accurate as possible so more people will use it. When colleagues denigrate my participation, my response is "If you find something inaccurate on Wikipedia, then you have an obligation to correct it!" I can't imagine how this could be more important than on a health-related page.*

Participants also felt responsible ensuring that the information was reliable and that people adhered to Wikipedia's editing guidelines. A number of people expressed concern about whether edits maintained an NPOV when concerning controversial topics. Exposed were also instances of when Wikipedia was used as a platform to promote an idea far beyond acceptance in the scientific community:

> *For vandalism and tendentious editing, my motivation is to maintain the integrity of Wikipedia. Vandalism irks and annoys me, but tendentious editing tends to infuriate me.*

> *...it makes me worried that instead of verifiability and notability being the driving things, it will be other agendas being pushed and that's disturbing. But I don't think Wikipedia is going away so I edit it. I continue to edit it because I think it must exist, it is important. And I just wish more people would edit it.*

Attitude Toward Wikipedia

Participant believed in the importance of building health-related content on Wikipedia:

> *Wikipedia is a beautiful, noble concept.*

The interviewed sample attributed strong positive beliefs to Wikipedia, which was identified as a strong motivating factor:

> *I use it. I support it. It's a thing worth doing. Wikipedia is a creation of lasting value.*

Thus, Wikipedian participants' beliefs and attitudes about Wikipedia were recognized as influencing their motivations and essential for the process of building Wikipedia. Despite the varying degrees of editorial skills and training of the participants, it became apparent that the interrelated motivations, which can be viewed as value systems, resulted in a motivated behavior model (see Figure 1).

Personal Fulfillment

Participants derived varying personal benefits in the process of editing health-related pages. Most proclaimed Wikipedia editing as being a hobby; others described it as fun, relaxing, engaging, and rewarding

> *I find it therapeutic. Yesterday I ended [up] writing an article rather than doing other things which are at the top of my to-do list.*

> *I felt slightly congratulated...so I think such things could be a positive motivator, seeing your work recognized in some form or description.*

Negative Experiences

Dedicated contributors reported engaging in debates or discussions either within their collaborative Wikipedia group (known as Wiki Projects) or over the article's discussion Talk pages, where both the contributors and readers frequently share experiences.

Their experiences varied but a proportion of the sample expressed being met with hostility and that was particularly relevant to controversial topics:

> *...it actually happened 1 or 2 times that it was quite aggressive and I lost my motivation for dealing with Wikipedia altogether.*

A number of interviewees expressed that they were not welcomed to the Wikipedia community despite their genuine intentions to contribute:

XSL•FO
RenderX

> *I had a bad experience with Wikipedia...but there is a culture associated with it in that the information has to be added in a certain way. So in good faith I added the content, it wasn't appropriate and there's a bit of rudeness on Wikipedia. People were rude to me about what I was trying to do because I did it incompetently. Looking back, I know what I did wrong and I did something inappropriate that was against the community rules but I was discouraged from contributing regularly at that time.*

This was also expressed in relation to how Wikipedia accounts are used. There was a divide in opinions about using accounts anonymously. Some emphasized openness because it made it easier for them to trust one another:

> *When people are hiding behind anonymity, they become a lot less nice. And on Wikipedia we already have a significant issue with civility problems.*

However, others expressed that anonymity is necessary to maintain the integrity of Wikipedia:

> *If I use my authority, then if my edits were wrong they might be accepted because I am in a position of authority, and that would be the opposite of a meritocracy. I think truth is the thing that should trump everything else, which means that authority has no place.*

Discussion

Value System and Intrinsic Motivation

Results were largely congruent with previous studies on Wikipedia in general, with some notable deviations. As in other studies [20-22,50] the vast majority of our participants were male, in full-time employment, and had obtained a university degree. However, we found a higher mean age and a substantially higher proportion of professionals (PhD, MD) compared to general studies of Wikipedians. We found broadly equal proportions of health specialists and laypeople in our sample.

To explain the underlying motivational drives, we found a process of interlinked value systems, compatible with the results of Kuznetsov [35]. There were 3 overarching primary motivation categories common to all contributors. These were expressed as education, which merged with the responsibility for maintaining accuracy (help), which merged with a sense of obligation (responsibility). These 3 overarching categories of motivations can be understood as self-efficacy as in Social Cognitive Theory [51] and support Kankanhalli and colleagues' [25] findings in which self-efficacy was found to be the most important predictor of knowledge-sharing behavior in online repositories. Furthermore, the primary overarching motivations were reported as inherent drives of the self, akin to the internal self-concept motivations proposed by Leonard et al [36]. This implies that Wikipedians are motivated to share knowledge because the process resonates with their internal values and beliefs.

It became apparent that motivated behavior arose not only from one's perception of the self, but also from the underlying beliefs about Wikipedia. Previous studies confirm that the degree of knowledge-sharing behavior significantly relates to individuals' perceptions of the context in which knowledge is shared [21,37]. Models such as the theory of reasoned action, the theory of planned behavior, or the technology acceptance model that are used to explain individual behavior equally recognize that attitude is crucial in knowledge-sharing intention and behavior [26,38,52].

A recent survey of Wikipedians indicated that 69% contribute to Wikipedia because of the ideology and 60% because they think it is fun [22] supporting the notion that positive outcomes are significant predictors of knowledge-sharing behavior [27] and supporting personal fulfillment as last of the emergent categories.

This study recognized that Wikipedians' differing editorial roles that can be understood as different levels in terms of "figure/ground" organization [52]. In other words, the ubiquitous motivation help was recognized as the processes of building and maintaining content expressed through various editorial activities of equal importance. This poses a challenge to the current view of Wikipedia in terms of knowledge sharing because not all Wikipedians engaged in knowledge sharing but instead in maintenance activity.

Implications

The aim of this study was to describe the characteristics of Wikipedians who edited health-related pages on Wikipedia and to gain an understanding of what drives them to contribute to Wikipedia's health-related pages. This study was the first qualitative exploration of Wikipedia contributors' motivations, not just in the health context but overall. With the exception of Yang and Lai [21,30], no integrated motivational model has been proposed to explain volunteer contributions in context of Wikipedia. Through the inductive process of grounded theory, no prior theoretical frameworks were "forced" on the data, allowing for the emergence of a realistic depiction of a concept directly from the data. The grounded theory analysis mapped a social process (depicted by Figure 1) of contributing knowledge on Wikipedia driven by individual's interrelated value systems.

The method of grounded theory also revealed an additional finding, hostility, with a possible connection to Wikipedia accounts' anonymity (a divide between the editors whose contributions are anonymous and those who, in part or in full, disclose their identity). Although these findings are suggestive rather than definite, they raise a challenge of whether Wikipedia's philosophy of equality is directly linked to anonymity. Wikipedia is egalitarian, a place where everybody are peers and a place where everyone has an equal right to edit contents. According to our results, a portion of Wikipedians believe that nonanonymous accounts would aid in civility, but also that nonanonymous accounts may create a hierarchy, a structure contradictory to Wikipedia's egalitarian philosophy.

This study also provided new evidence regarding the contributory behavior of Wikipedians: participants engage in contribution by utilizing their skill and not necessarily through knowledge sharing. Recently, a term was coined which describes Wikipedia contributors as "knowledge philanthropists" [53].

XSL•FO
RenderX

Although this term applied to a proportion of participants in this study, it is not applicable to all, particularly those who do not contribute to but instead "maintain" Wikipedia's content. Our broader view serves to recognize that everyone can contribute to Wikipedia without necessarily requiring expert knowledge.

Limitations

The sample of 44 articles used in the study may not be a representative sample of all health-related articles available on Wikipedia. The articles were randomly sampled from a total of approximately 3000 keywords complied from 3 medical databases and Portal Medicine's Featured Articles. An alternative approach would be to manually compile a list from Wikipedia's Category:Health, but the list would still not include all biomedical and drug-related articles.

Sampling bias may also apply to the recruitment of contributors. Selecting the most recent 5 contributors posed issues because some users appeared in the most recent 5 in more than 1 sampled article. In these instances, the researcher skipped accounts already contacted and contacted the next account down the list. This suggests that the editorial population of health content on Wikipedia is small. Another approach would be to select contributors according to the number of edits performed, although this may prove difficult because the numbers of edits are not necessarily indicative of editor's activity or the type of editorial involvement.

The response rate for the questionnaire was relatively low, for which the reasons may have been the mode and duration of the advertisement of the study. Only 32 participants completed the survey and 17 were interviewed. This is only a sample and does not represent all Wikipedians active on health-related articles. (We note the list of participants in WikiProject Medicine is much larger with 424 members as of August 2014 [54].) We suspect that this is a reflection of recruiting people via their Wikipedia user pages, which means participants had to be active on Wikipedia during the limited study period to see the recruitment message. It is fair to assume that the identified motivations might be sufficiently pervasive to be represented in a small sample of Wikipedians; however, varying levels of editorial skill and knowledge are not likely to be sampled deeply enough to be representative. The sample were recruited in a specific time frame and results may not be applicable over time. There are currently still challenges with increasing participation in contributing to Wikipedia health-related content. Some initiatives are already in place, such as the Translation Task Force and Wiki Project Med Foundation, a Wikipedia education program designed to educate medical students about the process and value of contribution to Wikipedia health pages, as well as also collaborating with a number of organizations including the Cochrane Collaboration, Cancer Research UK, and the National Institute of Health [55].

The success will largely depend on user's satisfaction and recognition of the potential benefit that can be gained from such editorial activities. By understanding Wikipedians' motivations for editing health-related content, we can better recruit more people to the task. Equally important is recognizing the factors that may discourage people and more specifically professionals, from contributing to Wikipedia. Characterizing editing behavior and editors also allows us to understand the processes underlying Wikipedia's health-related content.

Acknowledgments

The authors would like to thank all the Wikipedians who volunteered to participate in this study, Jenna Spink and Emma Green for double coding, and Benjamin Garner Sood for comments.

Conflicts of Interest

None declared. NF is on Wikipedia as Hydra Rain and HP is on as Bondegezou.

References

1. Wikipedia. 2012. Wikipedia: size of Wikipedia URL: http://en.wikipedia.org/wiki/Wikipedia:Size_of_Wikipedia [accessed 2014-06-14] [WebCite Cache ID 6QK2DuMv6]
2. Alexa. 2014. Alexa top 500 global sites URL: http://www.alexa.com/topsites [accessed 2014-06-14] [WebCite Cache ID 6QK2JqWJ0]
3. Wikipedia. 2012. Wikipedia: Wikipedians URL: http://en.wikipedia.org/wiki/Wikipedia:Wikipedians [accessed 2014-06-14] [WebCite Cache ID 6QK2W6xaA]
4. Wikimedia. 2012. User:Stu/comScore data on Wikimedia URL: http://meta.wikimedia.org/wiki/User:Stu/comScore_data_on_Wikimedia [accessed 2014-06-14] [WebCite Cache ID 6QK2eAzrn]
5. Powell J, Inglis N, Ronnie J, Large S. The characteristics and motivations of online health information seekers: cross-sectional survey and qualitative interview study. J Med Internet Res 2011;13(1):e20 [FREE Full text] [doi: 10.2196/jmir.1600] [Medline: 21345783]
6. Lustria ML, Smith SA, Hinnant CC. Exploring digital divides: an examination of eHealth technology use in health information seeking, communication and personal health information management in the USA. Health Informatics J 2011 Sep;17(3):224-243. [doi: 10.1177/1460458211414843] [Medline: 21937464]

XSL•FO
RenderX

7. Kummervold PE, Chronaki CE, Lausen B, Prokosch HU, Rasmussen J, Santana S, et al. eHealth trends in Europe 2005-2007: a population-based survey. J Med Internet Res 2008;10(4):e42 [FREE Full text] [doi: 10.2196/jmir.1023] [Medline: 19017584]
8. Lewis T, Synowiec C, Lagomarsino G, Schweitzer J. E-health in low- and middle-income countries: findings from the Center for Health Market Innovations. Bull World Health Organ 2012 May 1;90(5):332-340 [FREE Full text] [doi: 10.2471/BLT.11.099820] [Medline: 22589566]
9. Health Fact Sheet. Washington, DC: Pew Internet & American Life Project; 2012. URL: http://www.pewinternet.org/fact-sheets/health-fact-sheet/ [accessed 2014-06-14] [WebCite Cache ID 6QK2hsmuD]
10. Dutton WH, Helsper EJ, Gerber MM. The Internet in Britain 2009. Oxford: Oxford Internet Institute; 2009. Oxford Internet Survey URL: http://oxis.oii.ox.ac.uk/ [accessed 2014-06-16] [WebCite Cache ID 6QNJgIV40]
11. Moretti FA, Oliveira VE, Silva EM. Access to health information on the internet: a public health issue? Rev Assoc Med Bras 2012;58(6):650-658 [FREE Full text] [Medline: 23250092]
12. Wikipedia. 2012. Wikipedia:WikiProject Medicine/Popular pages URL: http://en.wikipedia.org/wiki/Wikipedia:WikiProject_Medicine/Popular_pages [WebCite Cache ID 6QK21762EuI]
13. Laurent MR, Vickers TJ. Seeking health information online: does Wikipedia matter? J Am Med Inform Assoc 2009 Apr;16(4):471-479 [FREE Full text] [doi: 10.1197/jamia.M3059] [Medline: 19390105]
14. Archambault PM, van de Belt TH, Grajales Iii FJ, Eysenbach G, Aubin K, Gold I, et al. Wikis and collaborative writing applications in health care: a scoping review protocol. JMIR Res Protoc 2012;1(1):e1 [FREE Full text] [doi: 10.2196/resprot.1993] [Medline: 23612481]
15. Waters NL. Why you can't cite Wikipedia in my class. Commun ACM 2007 Sep 01;50(9):15. [doi: 10.1145/1284621.1284635]
16. Giles J. Internet encyclopaedias go head to head. Nature 2005 Dec 15;438(7070):900-901. [doi: 10.1038/438900a] [Medline: 16355180]
17. Czarnecka-Kujawa K, Abdalian R, Grover SC. The quality of open access and open source Internet material in gastroenterology: is Wikipedia appropriate for knowledge transfer to patients? Gastroenterology 2008 Apr;134(4):A-325-A-326. [doi: 10.1016/s0016-5085(08)61518-8]
18. Heilman JM, Kemmann E, Bonert M, Chatterjee A, Ragar B, Beards GM, et al. Wikipedia: a key tool for global public health promotion. J Med Internet Res 2011;13(1):e14 [FREE Full text] [doi: 10.2196/jmir.1589] [Medline: 21282098]
19. West R, Weber I, Castillo C. Drawing a Data-Driven Portrait of Wikipedia Editors. 2012. URL: http://infolab.stanford.edu/~west1/pubs/wikiedits.pdf [accessed 2014-06-14] [WebCite Cache ID 6QK2uui3D]
20. Schroer J, Hertel G. Voluntary engagement in an open Web-based encyclopedia: Wikipedians and why they do it. Media Psychology 2009 Mar 05;12(1):96-120. [doi: 10.1080/15213260802669466]
21. Yang H, Lai C. Understanding knowledge-sharing behaviour in Wikipedia. Behaviour & Information Technology 2011 Jan;30(1):131-142. [doi: 10.1080/0144929X.2010.516019]
22. Wikimedia. 2012. Editor Survey 2011 URL: http://meta.wikimedia.org/wiki/Editor_Survey_2011 [accessed 2014-11-04] [WebCite Cache ID 6TqFp56US]
23. Pfeil U, Zaphiris P, Ang CS. Cultural differences in collaborative authoring of Wikipedia. J Comp Mediated Comm 2006 Oct;12(1):88-113. [doi: 10.1111/j.1083-6101.2006.00316.x]
24. Amichai-Hamburger Y, Lamdan N, Madiel R, Hayat T. Personality characteristics of Wikipedia members. Cyberpsychol Behav 2008 Dec;11(6):679-681. [doi: 10.1089/cpb.2007.0225] [Medline: 18954273]
25. Kankanhalli A, Tan BCJ, Wei KK. Contributing knowledge to electronic knowledge repositories: An empirical investigation. MIS Quarterly 2005 Mar;29(1):113-143 [FREE Full text]
26. Bock GW, Zmund RW, Kim YG, Lee JN. Behavioural intention formation in knowledge sharing: Examining the roles of extrinsic motivators, social-psychological forces, and organizational climate. MIS Quart 2005 Mar;29(1):87-111 [FREE Full text]
27. Cabrera A, Collins WC, Salgado JF. Determinants of individual engagement in knowledge sharing. The International Journal of Human Resource Management 2006 Feb;17(2):245-264. [doi: 10.1080/09585190500404614]
28. Wasko MM, Faraj S. Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. MIS Quarterly 2005 Mar;29(1):35-57 [FREE Full text]
29. Deci EL. Intrinsic Motivation. New York: Plenum Publishing Co; 1975.
30. Yang HL, Lai CY. Motivations of Wikipedia content contributors. Comput Hum Behav 2010;26:1377. [doi: 10.1016/j.chb.2010.04.011]
31. Benkler Y, Nissenbaum H. Commons-based peer production and virtue. J Pol Philo 2006;14(4):394-419 [FREE Full text]
32. Jenkins H. Convergence Culture: Where Old and New Media Collide. New York: New York University Press; 2006.
33. Bryant SL, Forte A, Bruckman A. Becoming Wikipedian: transformation of participation in a collaborative online encyclopaedia. In: GROUP '05 Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work. 2005 Presented at: GROUP '05 International Conference on Supporting Group Work; November 6-9, 2005; Sanibel Island, FL p. 1-10. [doi: 10.1145/1099203.1099205]
34. Royal C, Kapila D. What's on Wikipedia, and What's Not...?: Assessing Completeness of Information. Social Science Computer Review 2008 Apr 07;27(1):138-148. [doi: 10.1177/0894439308321890]

XSL•FO
RenderX

35. Kuznetsov S. Motivations of contributors to Wikipedia. SIGCAS Comput Soc 2006 Jun 01;36(2):1-es. [doi: 10.1145/1215942.1215943]
36. Leonard NH, Beauvais LL, Scholl RW. Work motivation: the incorporation of self-concept based processes. Hum Relat 1999;52(8):969-998 [FREE Full text]
37. Lin HF. Effects of extrinsic and intrinsic motivation on employee knowledge sharing intentions. J Inf Sci 2007;33(2):135-149. [doi: 10.1177/0165551506068174]
38. Bock GW, Kim YG. Breaking the myths of rewards: An explanatory study of attitudes and validation of new scales to measure an integrative taxonomy of motivation. Psychol Rep 2002;82(3):1011-1022. [doi: 10.4018/irmj.2002040102]
39. Maslow AH. Motivation and Personality. New York: Harper and Row; 1987.
40. Wikipedia. Portal:Medicine/Selected Article URL: http://en.wikipedia.org/?title=Portal:Medicine/Selected_Article [accessed 2014-11-02] [WebCite Cache ID 6TnJqVTYG]
41. Wikipedia. User:Hydra Rain URL: http://en.wikipedia.org/wiki/User:Hydra_Rain [accessed 2014-11-02] [WebCite Cache ID 6TnKw01Iy]
42. Opinio Survey Software. A study of editors of health-related Wikipedia pages URL: https://opinio.ucl.ac.uk/s?s=18871 [accessed 2014-11-02] [WebCite Cache ID 6TnL0BXhq]
43. Elliot R, Fischer CT, Rennie DL. Evolving guidelines for publication of qualitative research studies in psychology and related fields. Brit J Clin Psychol 1999;38:215-229. [doi: 10.1348/014466599162782]
44. Glaser BG, Strauss AL. The Discovery of Grounded Theory: Strategies for Qualitative Research. Chicago, IL: Aldine Pub Co; 2012.
45. Willig C. Introducing qualitative research in psychology: adventures in theory and method. Maidenhead: Open University Press/McGraw-Hill; 2008.
46. Nov O. What motivates Wikipedians? Commun ACM 2007 Nov 01;50(11):60-64. [doi: 10.1145/1297797.1297798]
47. Wikipedia. Wikipedia:Verifiability URL: http://en.wikipedia.org/wiki/Wikipedia:Verifiability [accessed 2014-11-02] [WebCite Cache ID 6TnLECPEj]
48. Wikipedia. Wikipedia:Neutral point of view URL: http://en.wikipedia.org/wiki/Wikipedia:Neutral_point_of_view [accessed 2014-11-05] [WebCite Cache ID 6TrRnfGpG]
49. Wikimedia. Wikipedia:Identifying reliable sources (medicine) URL: http://en.wikipedia.org/wiki/Wikipedia:Identifying_reliable_sources_(medicine) [accessed 2014-11-02] [WebCite Cache ID 6TnLmwe9E]
50. Bandura A, Adams NE, Beyer J. Cognitive processes mediating behavioral change. J Pers Soc Psychol 1977 Mar;35(3):125-139. [Medline: 15093]
51. Ajzen I. The theory of planned behavior. Organizational Behavior and Human Decision Processes 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
52. Perls FS, Hefferline RF, Goodman P. Gestalt Therapy: Excitement and Growth in the Human Personality. London: Souvenir Press; 1972.
53. Bastian H. Scientific American. 2012 Jul 18. Are You a Knowledge Philanthropist? If Not, Why Not? URL: http://blogs.scientificamerican.com/guest-blog/2012/07/18/are-you-a-knowledge-philanthropist-if-not-why-not/ [accessed 2014-06-14] [WebCite Cache ID 6QK39GkzL]
54. Wikipedia. Wikipedia:WikiProject Medicine/Participants URL: http://en.wikipedia.org/wiki/Wikipedia:WikiProject_Medicine/Participants [accessed 2014-08-17] [WebCite Cache ID 6RtkCdRSs]
55. Wikimedia. Wiki Project Med URL: https://meta.wikimedia.org/wiki/WikiProject_Med [accessed 2014-11-02] [WebCite Cache ID 6TnMQcGIh]

Abbreviations

IP: Internet Protocol
NHS: National Health Service
NORD: National Organization of Rare Disorders
NPOV: neutral point of view

Edited by G Eysenbach; submitted 18.06.14; peer-reviewed by M Laurent, J Heilman, T van de Belt; comments to author 17.07.14; revised version received 04.08.14; accepted 20.08.14; published 03.12.14.

Please cite as:
Farič N, Potts HWW
Motivations for Contributing to Health-Related Articles on Wikipedia: An Interview Study
J Med Internet Res 2014;16(12):e260
URL: http://www.jmir.org/2014/12/e260/
doi:10.2196/jmir.3569
PMID:25498308

XSL•FO
RenderX

©Nuša Farič, Henry WW Potts. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 03.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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

Analysis of a Parent-Initiated Social Media Campaign for Hirschsprung's Disease

Kristy Wittmeier[1,2,3], PT, MSc, PhD; Cindy Holland[4], RN, MN, NP; Kendall Hobbs-Murison[3,4,5], RN, BN; Elizabeth Crawford[6]; Chad Beauchamp[7], MA, Dm(hons); Brodie Milne[7]; Melanie Morris[2,3,5], MD, FRCSC; Richard Keijzer[2,3,5], MSc, MD, PhD

[1]Centre for Healthcare Innovation, University of Manitoba, Winnipeg Regional Health Authority, Winnipeg, MB, Canada

[2]Department of Pediatrics and Child Health, University of Manitoba, Winnipeg, MB, Canada

[3]Manitoba Institute of Child Health, Winnipeg, MB, Canada

[4]Child Health Program, Health Sciences Centre, Winnipeg, MB, Canada

[5]Department of Surgery, University of Manitoba, Winnipeg, MB, Canada

[6]Swish Productions Ltd., Winnipeg, MB, Canada

[7]Direct Focus Marketing Communications Inc., Winnipeg, MB, Canada

Corresponding Author:
Richard Keijzer, MSc, MD, PhD
Department of Surgery
University of Manitoba
AE 402
840 Sherbrook Street
Winnipeg, MB, R3A 1S1
Canada
Phone: 1 204 787 8854
Fax: 1 204 787 1479
Email: richardkeijzer@gmail.com

Abstract

Background: Social media can be particularly useful for patients or families affected by rare conditions by allowing individuals to form online communities across the world.

Objective: Our aim in this study was to conduct a descriptive and quantitative analysis of the use of a social media community for Hirschsprung's Disease (HD).

Methods: In July 2011, a mother of a child with HD launched the "Shit Happens" campaign. The campaign uses social media (blogs, Twitter, and Facebook) to engage other families affected by HD. Internet analytics including Google Analytics and Facebook Insights were used to evaluate the reach and responsiveness of this campaign.

Results: On the day the HD campaign was launched, 387 people viewed the blog "Roo's Journey". Blog views have now exceeded 5400 views from 37 countries. The Facebook page extends to 46 countries, has an average post reach of 298 users, 1414 "likes", and an overall reach of 131,032 users. The campaign has 135 Twitter followers and 344 tweets at the time of writing. The most common question posted on the Facebook page is related to treatment for extreme diaper rash. Responsiveness assessment demonstrated that within 2 hours of posting, a question could receive 143 views and 20 responses, increasing to 30 responses after 5 hours.

Conclusions: Social media networks are well suited to discussion, support, and advocacy for health-related conditions and can be especially important in connecting families affected by rare conditions. The HD campaign demonstrates the reach and responsiveness of a community that primarily relies on social media to connect families affected by HD. Although responsive, this community is currently lacking consistent access to evidence-based guidance for their common concerns. We will explore innovative consumer-researcher partnerships to offer a solution in future research.

(J Med Internet Res 2014;16(12):e288) doi:10.2196/jmir.3200

XSL•FO
RenderX

KEYWORDS

Hirschsprung's Disease; social media; patient oriented research; knowledge translation

Introduction

Hirschsprung's disease (HD) is a rare condition that is estimated to affect 1 in 5000 live births [1]. This disease primarily affects the innervation of the gastrointestinal tract due to failure in neural crest cell migration during the embryonic period. As a result, a section of bowel is left aganglionic, resulting in a portion of bowel essentially paralyzed and unable to mobilize stool [2].

HD typically presents in infancy with a delay or failure to pass meconium, ongoing difficulty passing stool, abdominal distension, and intolerance to feeding. However, HD may also present in early childhood or later with mild symptoms in association with growth concerns and chronic constipation. Treatment generally consists of a pull-through surgery, where the affected segment of bowel is removed and the healthy portion of bowel is re-anastomosed. Historically, there has been advancement from a three-step procedure to current practice of a single or two-step procedure with use of minimally invasive techniques when appropriate [3]. Researchers continue to investigate the possible use of cell transplantation in the antenatal period [3,4].

Despite promising advances in diagnostics and surgical techniques, there remain many challenges that individuals with HD and their caregivers encounter in the day-to-day management of this condition. In the postoperative period, challenges include irregular stool patterns, ongoing incontinence, impaired psychosocial function, the potential of enterocolitis, and concerns regarding quality of life for both the child and caregiver [4,5].

Specific to caregivers, it is becoming better recognized that raising a child with a rare congenital anomaly presents a number of unique challenges that can contribute to a higher level of parental stress [6]. Stress is associated with concerns regarding the competencies required to care for a child with medical needs, or feelings of social isolation [7,8]. This is exaggerated in large countries such as Canada and the United States where the closest pediatric hospital might be hundreds of kilometers away, making it difficult to obtain relevant advice or adequate support in a timely fashion. Traditional social support networks such as family, friends, and neighbors may be inadequate or ill equipped to understand and address the issues and questions specific to parenting a child with complex needs, which further aggravates parental stress.

The Internet is a common resource accessed by the general public on medical issues [9]. A recent report indicates that 59% of adult Internet users search for health information online [9]. Simply searching "Hirschsprung's Disease" on Google returns over 660,000 results within 0.24 seconds, with the top hits containing detailed information on the disease definition, diagnosis, and surgical procedures used to manage the condition. Information to guide caregivers in day-to-day intricacies of the life of a child with HD as well as on longer-term complications of the condition is unfortunately less available to the general public.

Social media overcomes the limitations of more static websites designed primarily to disseminate information, by facilitating the development of relationships and communities around shared interests [10,11]. Social media provides a portal to quick, interactive communication and a venue for people with common interests, including health-related interests to connect [9], such as those affected by chronic and rare diseases [12-14]. Acknowledging that social media has changed the way people collect and disseminate information [15], the potential for social media to help improve disease management and health outcomes requires further study.

In an effort to build a visible and easily accessible interactive Web-based community for caregivers of children with HD, a parent within the HD community in Manitoba, Canada, has collaborated with a marketing team to develop, implement, and maintain an ongoing social media campaign, entitled "Shit Happens". This mother (EC; "Liz") has partnered with our research team to evaluate the campaign as part of a broader effort to understand and improve the care of children with HD nationally and internationally. The primary objective of this study was to perform a descriptive analysis of the HD social media campaign and to conduct a quantitative assessment of the reach and responsiveness of the community. A secondary objective was to conduct a preliminary content analysis of priority areas for health information needs as identified by HD social media community members.

Methods

Overview

The targeted audience of this campaign was caregivers of children with HD and those living with HD that self-engaged in social media. Health care professionals were not a targeted group at this time. The campaign has facilitated the development of a virtual community, defined as "online social networks in which people with common interests, goals, or practices interact to share information and knowledge, and engage in social interactions" [16]. The community and their interactions were each analyzed through a combination of embedded performance measures and available Internet analysis tools on the blog, Facebook, and Twitter. Collectively, these methods provide data to assess the reach and responsiveness of the HD social media campaign, indicators of community engagement, and campaign impact.

Reach

"The Roo's Journey" blog was analyzed using Google Analytics allowing real-time and custom data reporting through the development of an indicator dashboard. Both dimensions (describe data) and metrics (measure data) were built in to provide a custom report [17]. Specific items monitored on the dashboard for Roo's Journey include new visits, unique visitors, visits by (Internet) browser, mapping (geographical

XSL•FO
RenderX

representation of visits), average visit duration, and pages per visit. Facebook's Insights dashboard provides data on the number of countries the account has reached, age category and gender of users, potential reach (includes friends of page fans), "likes", post reach (number of unique people who saw that individual post), and feedback. Past data provided include number of active monthly users and total post views. Twitter analytics include the number of followers and number of tweets.

Responsiveness

Community responsiveness was assessed by quantifying the volume and timing of feedback for a question posted to Facebook by the site administrator. Priority areas for information needs were determined through a qualitative analysis and identification of recurring themes within questions posed by the Facebook community members. As the social media platforms were within the public domain and the data were collected by the site administrator, the University of Manitoba Research Ethics Office waived the need for ethical review of this study.

Results

Overview

In July 2011, the mother of a 2-year old with HD, launched the "Shit Happens" campaign (Figure 1).

This campaign included a blog titled Roo's Journey, describing a family's experience with HD, a Facebook page (Hirschsprung's Community), and a Twitter account (@shithapps). The collective objective of these three forums was to support families living with HD and raise awareness of the disease. The HD campaign name "Shit Happens" was chosen from a marketing perspective to create a brand that would stand out to enhance campaign uptake and spread. The campaign name was also chosen to resonate with potential community members and break the barriers of an otherwise socially awkward topic. In Liz's words:

> *Many children with HD suffer from chronic constipation and occasional bowel obstruction. For a child and family with HD the idea of "shit" happening tends to bring feelings of joy and relief. When a bowel movement finally occurs it is not for the faint of heart…hence the name (Shit Happens). Laugh or cry…may as well laugh!*

The campaign creator is the primary contributor and chief administrator of all three social media platforms. The Hirschsprung's community on Facebook has daily interaction as questions and concerns arise. While the Facebook page was created to engage caregivers, its primary function was to create a support network. Online and offline supportive relationships and networks have developed through introductions and referrals. Facebook posts were shared as tweets through Twitter.

The Roo's Journey blog provides an in-depth and honest portrayal of caring for a child with HD. Facebook and Twitter are used to share links to Roo's Journey, other blogs, videos, and information on conferences and events that may be of interest to the HD community. In 2012, the HD Around the Globe campaign was launched within the HD Community using all social media platforms to obtain wider recognition of the disease. This ongoing campaign features photos of individuals from the HD community, celebrities, and models wearing the HD Community "SHIT HAPPENS" T-shirt (Figure 2).

To ensure the sites are efficiently managed, the posts are monitored daily. All Facebook posts are screened prior to reposting, and posts that are considered inappropriate or potentially harmful are immediately deleted. Since inception, there have been only two post deletions; these were due to misuse of language related to expression of frustration. All posts and comments to these sites can be viewed publicly, while personal messaging or email to the creator in the case of the blog remain private. For Facebook and Twitter, participating users must agree to the terms and conditions outlined by these social media platforms. HD community participation requires no additional inclusion criteria, although it is noted that the groups were created as a resource to those affected by HD. Outside of the social media realm, this campaign extends promotion of the Hirschsprung's Community through events such as fashion events in shopping centres, Fashion Rocks (an annual fundraiser), and Easton's memorial poker run (Nebraska, USA). The campaign creator's personal public relations company sustains the campaign in collaboration with an affiliated marketing company (Swish Productions Ltd. and Direct Focus Marketing Communications Inc.). Local and international health care professionals with expertise in HD demonstrate campaign support with presentations and participation at HD fundraising opportunities and by informing families whom they interact with about this ongoing campaign.

The most common question posted on the HD Facebook page is regarding the management of severe diaper rash, followed closely by questions regarding bowel treatments, long-term outcomes, potty training, and dietary issues (Figure 3).

XSL•FO
RenderX

Figure 1. Blog, Facebook, and Twitter logos for HD campaign.

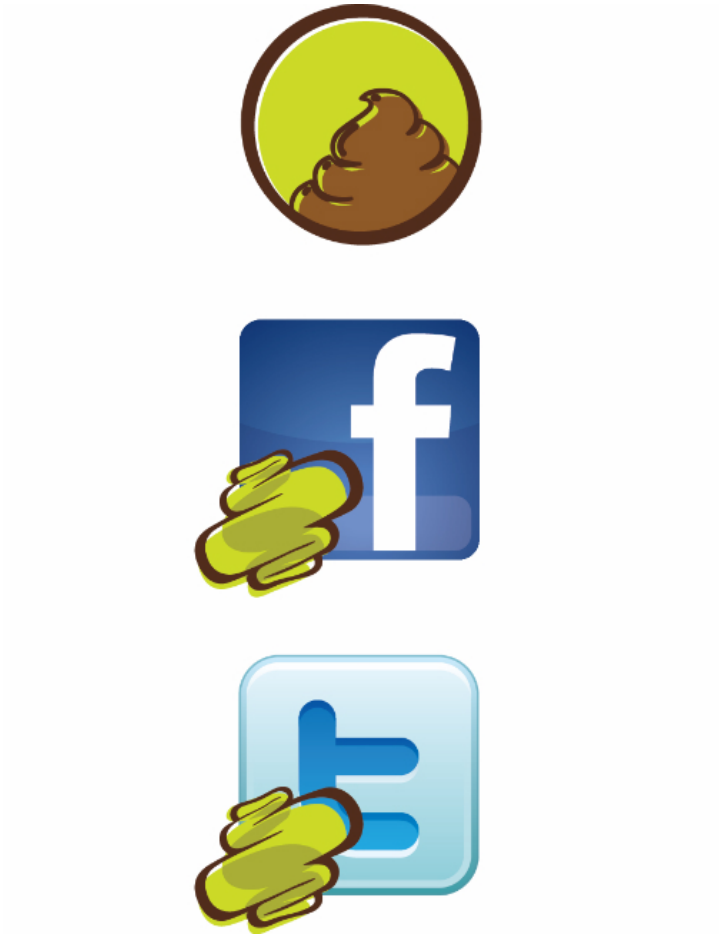
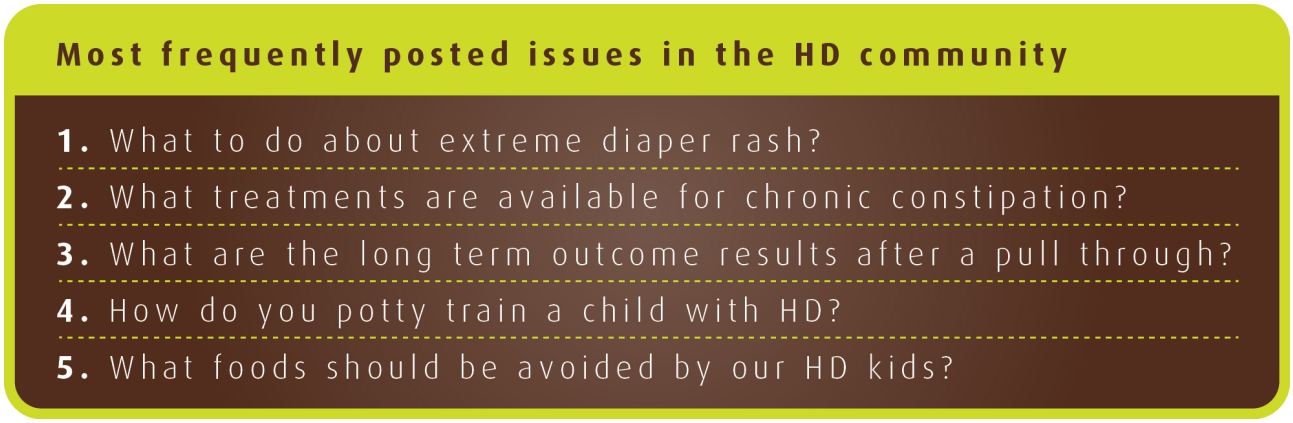


Figure 2. Image from HD Around the Globe campaign featuring models wearing the campaign T-shirt.



XSL•FO
RenderX

Figure 3. The five most common issues raised in HD community on the Facebook page.



Reach

On the day of the launch of the HD Community Campaign, 387 people viewed the Roo's Journey blog. Almost 3 years post-launch, there have been 5487 blog views and/or visitors, encompassing 37 countries. The blog has been most often read in the United States, Canada, and the United Kingdom.

The Facebook page extends to 46 countries with the average user being female between the ages of 25-44 years (Figure 4).

Past metrics provided by Facebook Insights indicate that there were 1130 monthly users of the HD Community Facebook page. At the time of writing, there were a total of 1414 "likes",and based on the last 28 days, the average post reach (number of unique people who viewed the post) was 298 people, with a total overall reach of the community (includes Facebook page users and their friends) of 131,032 people.

On Twitter there are currently 135 followers, and 344 tweets have been posted.

Figure 4. Reach of HD Community Facebook page.



Responsiveness

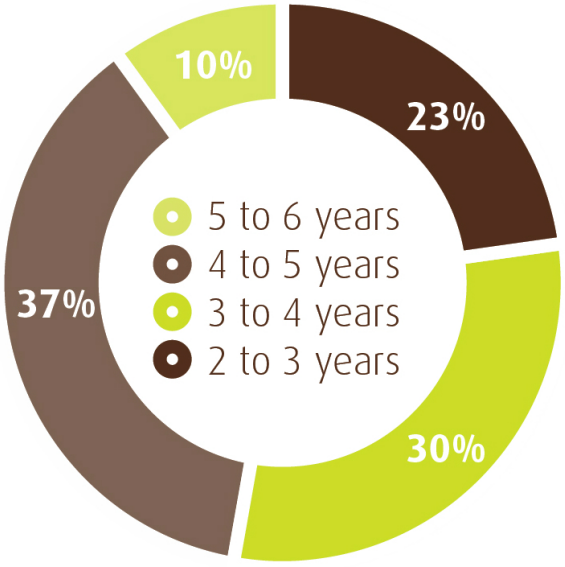
Community responsiveness was measured with a post asking, "At what age was your child potty trained?". The question was viewed 143 times within 2 hours, and 20 individuals provided responses (14% of viewers). Five hours after the question was posed, 30 responses were posted (Figure5) with the majority indicating that despite being potty trained, their children still experienced ongoing incontinence issues.

Similarly, when a question was posted regarding management of diaper rash and diarrhea after pull-through surgery, 413

XSL•FO
RenderX

people viewed the post within 2 hours and 20 members (5% of viewers) replied offering advice and support on how to deal with this issue based on their experience.

Figure 5. Reported percentages of "age of potty training" in response to the question "At what age was your child potty trained?".



Discussion

Principal Findings

Social media can be a powerful tool to create a community for caregivers of children born with rare diseases such as HD, connecting individuals across the world within minutes. Demonstrating reach of a large geography and many participants, this caregiver community is engaged and responsive. A content analysis of Facebook postings highlights current priority areas for knowledge and resources among caregivers within this community.

From a health research perspective, this community provides the ideal opportunity for integrated knowledge translation by providing a venue for interaction between researchers and knowledge users [18]. A caregiver identified a gap in available resources for parents and has created a community to address this gap. Community members now use the social media platforms as information sharing tools, to exchange knowledge, advice, and experiences. What has been notably absent however, is evidence-informed input from health care providers—a trend observed in previous evaluation of health-related social media forums [19]. In this age of evidence-based practice and knowledge translation, it is necessary to consider the role of the clinician and the researcher within the social media sphere. Now, through this new partnership between caregivers, academics, and clinicians, the HD community is well poised to systematically identify and address knowledge gaps with evidence-based tools and resources.

Moving forward, a detailed analysis of issues raised within the HD community will provide the basis of a needs assessment to inform new research including knowledge synthesis and translation activities. Previous research examining Twitter and smoking cessation has highlighted inconsistencies in advice or tweet content relative to existing treatment guidelines [19]. A researcher-knowledge user partnership can help overcome the risk of sharing misinformation associated with consumers seeking health care advice via the Internet.

Guidelines and ethics documents are becoming more widely available for clinicians and researchers who wish to engage in social media [15,20-23]. These should be reviewed and adhered to appropriately. Basic considerations include respecting patient confidentiality, maintaining appropriate boundaries, disclosure of competing or conflicting interests, and consideration of professional reputation [20-23]. The same standards apply to patient-caregiver interactions in person and online, and if respected, can safely broaden the reach of the medical and research community in an effort to provide evidence-based answers to the complex issues that families face.

Strengths and Limitations

This growing community to support families affected by HD demonstrates a level of interaction that suggests satisfaction and commitment of users, an interrelationship that has an effect on sustainment of a virtual community [24]. The campaign has been designed and marketed to catch attention and resonate with members to ensure viability and sustainability.

Limitations of this study include the retrospective nature of our analysis. As this is a parent-initiated and parent-led campaign, it was not designed at inception to be scientifically evaluated. Internet analytics however, in combination with the vigilant management of this community have provided detailed quantitative and qualitative data on a real-world campaign dedicated to a health issue. Second, although a geographically wide representation has been achieved, this study includes caregivers of children with HD who have Internet access, the knowledge and skills to access online forums, and the desire to participate in one of the three online forums. Results specifically pertaining to the content analysis may not be representative of the overall caregiver community. Finally, this study was not designed to assess the value of the HD Community to its members. Further research is needed to understand the effect

XSL•FO
RenderX

of these types of campaigns on individuals who view and actively participate in the conversations (active users) as well as those who view information but do not personally contribute (passive users).

Conclusions

Social media has changed the way people communicate. This study demonstrates that by using a multipronged social media strategy, one has the ability to break through barriers of distance (global) and time (instant publication) while targeting the desired demographic. These characteristics make social media an effective tool to build a community as well as collect and disseminate information. The Hirschsprung's community has been connected through the Shit Happens Campaign. An HD Twitter account, Facebook page, and a personal blog dedicated to supporting families living with HD have been viewed hundreds of thousands of times across the globe. The sites are presently filled with questions from families who continue to struggle to manage even after surgical intervention for HD, presenting health care professionals and researchers with the opportunity to provide families with evidence-based information to guide care. Families rely on these forums for support from other caregivers, but clinicians and researchers are not represented within these social media communities. Partnerships between communities of caregivers, health clinicians, and researchers mediated through social media could provide unprecedented opportunity for consumer-driven research. Together, families and care providers can ensure that caregiver concerns become research priorities and that existing evidence and research results are widely and appropriately shared.

Acknowledgments

We wish to acknowledge the funding source: the Thorlakson Chair in Surgical Research. The funding source had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Authors' Contributions

Kristy Wittmeier and Richard Keijzer had full access to all the data in the study and take responsibility for the integrity of the data and accuracy of the data analysis. Chad Beauchamp and Brodie Milne (Direct Focus Marketing Communications Inc.) managed the creative strategy and conducted the data analysis involving Internet analytics. Kristy Wittmeier, Cindy Holland, Kendall Hobbs-Murison, Melanie Morris, and Richard Keijzer analyzed the data. Elizabeth Crawford's company, Swish Productions Ltd., has been responsible for the campaign strategy and public relations relating to the campaign since inception. Kristy Wittmeier, Cindy Holland, Kendall Hobbs-Murison, Elizabeth Crawford, Chad Beauchamp, Brodie Milne, Melanie Morris, and Richard Keijzer wrote and revised the manuscript.

Conflicts of Interest

None declared.

References

1. Best KE, Glinianaia SV, Bythell M, Rankin J. Hirschsprung's disease in the North of England: prevalence, associated anomalies, and survival. Birth Defects Res A Clin Mol Teratol 2012 Jun;94(6):477-480. [doi: 10.1002/bdra.23016] [Medline: 22511583]
2. de Lorijn F, Boeckxstaens GE, Benninga MA. Symptomatology, pathophysiology, diagnostic work-up, and treatment of Hirschsprung disease in infancy and childhood. Curr Gastroenterol Rep 2007 Jun;9(3):245-253. [Medline: 17511924]
3. Gunnarsdóttir A, Wester T. Modern treatment of Hirschsprung's disease. Scand J Surg 2011;100(4):243-249. [Medline: 22182845]
4. Wilkinson DJ, Edgar DH, Kenny SE. Future therapies for Hirschsprung's disease. Semin Pediatr Surg 2012 Nov;21(4):364-370. [doi: 10.1053/j.sempedsurg.2012.07.010] [Medline: 22985842]
5. Hartman EE, Oort FJ, Aronson DC, Sprangers MA. Quality of life and disease-specific functioning of patients with anorectal malformations or Hirschsprung's disease: a review. Arch Dis Child 2011 Apr;96(4):398-406. [doi: 10.1136/adc.2007.118133] [Medline: 20371581]
6. Fonseca A, Nazaré B, Canavarro MC. Parental psychological distress and quality of life after a prenatal or postnatal diagnosis of congenital anomaly: a controlled comparison study with parents of healthy infants. Disabil Health J 2012 Apr;5(2):67-74. [doi: 10.1016/j.dhjo.2011.11.001] [Medline: 22429541]
7. Butcher PR, Wind T, Bouma A. Parenting stress in mothers and fathers of a child with a hemiparesis: sources of stress, intervening factors and long-term expressions of stress. Child Care Health Dev 2008 Jul;34(4):530-541. [doi: 10.1111/j.1365-2214.2008.00842.x] [Medline: 19154554]
8. Dellve L, Samuelsson L, Tallborn A, Fasth A, Hallberg LR. Stress and well-being among parents of children with rare diseases: a prospective intervention study. J Adv Nurs 2006 Feb;53(4):392-402. [doi: 10.1111/j.1365-2648.2006.03736.x] [Medline: 16448482]

XSL•FO
RenderX

9. Fox S, Duggan M. Health Online. 2013. URL: http://www.pewinternet.org/Reports/2013/Health-online/Summary-of-Findings.aspx [accessed 2013-12-11] [WebCite Cache ID 6LnAj58UR]
10. Sugawara Y, Narimatsu H, Hozawa A, Shao L, Otani K, Fukao A. Cancer patients on Twitter: a novel patient community on social media. BMC Res Notes 2012;5:699 [FREE Full text] [doi: 10.1186/1756-0500-5-699] [Medline: 23270426]
11. Thoren EM, Metze B, Bührer C, Garten L. Online support for parents of preterm infants: a qualitative and content analysis of Facebook 'preemie' groups. Arch Dis Child Fetal Neonatal Ed 2013 Nov;98(6):F534-F538. [doi: 10.1136/archdischild-2012-303572] [Medline: 23759520]
12. 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. [doi: 10.1016/j.jbi.2013.04.010] [Medline: 23702104]
13. Schumacher KR, Stringer KA, Donohue JE, Yu S, Shaver A, Caruthers RL, et al. Social media methods for studying rare diseases. Pediatrics 2014 May;133(5):e1345-e1353. [doi: 10.1542/peds.2013-2966] [Medline: 24733869]
14. Thackeray R, Crookston BT, West JH. Correlates of health-related social media use among adults. J Med Internet Res 2013;15(1):e21 [FREE Full text] [doi: 10.2196/jmir.2297] [Medline: 23367505]
15. O'Connor D. The apomediated world: regulating research when social media has changed research. J Law Med Ethics 2013;41(2):470-483. [doi: 10.1111/jlme.12056] [Medline: 23802898]
16. Chiu CM, Hsu MH, Wang ETG. Understanding knowledge sharing in virtual communities: An integration of social capital and social cognitive theories. Decision Support Systems 2006 Dec;42(3):1872-1888. [doi: 10.1016/J.Dss.2006.04.001]
17. Google. Google Analytics: Dimensions and metrics URL: https://support.google.com/analytics/answer/1033861?rd=1 [accessed 2013-12-10] [WebCite Cache ID 6LnB2Qyw4t]
18. CIHR. More About Knowledge Translation at CIHR. 2013. URL: http://www.cihr-irsc.gc.ca/e/39033.html [accessed 2013-12-11] [WebCite Cache ID 6LnBAHoq2]
19. Prochaska JJ, Pechmann C, Kim R, Leonhardt JM. Twitter=quitter? An analysis of Twitter quit smoking social networks. Tob Control 2012 Jul;21(4):447-449 [FREE Full text] [doi: 10.1136/tc.2010.042507] [Medline: 21730101]
20. McKee R. Ethical issues in using social media for health and health care research. Health Policy 2013 May;110(2-3):298-301. [doi: 10.1016/j.healthpol.2013.02.006] [Medline: 23477806]
21. American Medical Association. Report of the Council on Ethical and Judicial Affairs. Professionalism in the Use of Social Media URL: http://www.ama-assn.org/resources/doc/code-medical-ethics/9124a.pdf [accessed 2013-12-11] [WebCite Cache ID 6LnBEY63n]
22. St-Laurent-Gagnon T, Coughlin KW, Canadian Paediatric Society, Bioethics Committee. Paediatricians, social media and blogs: Ethical considerations. Paediatr Child Health 2012 May;17(5):267-272 [FREE Full text] [Medline: 23633902]
23. Bik HM, Goldstein MC. An introduction to social media for scientists. PLoS Biol 2013;11(4):e1001535 [FREE Full text] [doi: 10.1371/journal.pbio.1001535] [Medline: 23630451]
24. Cheung C, Lee M. Understanding the sustainability of a virtual community: model development and empirical test. Journal of Information Science 2009 Mar 24;35(3):279-298 [FREE Full text] [doi: 10.1177/0165551508099088]

Edited by G Eysenbach; submitted 03.01.14; peer-reviewed by A Goldstein, M Mazzucato; comments to author 28.04.14; revised version received 07.06.14; accepted 17.08.14; published 11.12.14.

Please cite as:
Wittmeier K, Holland C, Hobbs-Murison K, Crawford E, Beauchamp C, Milne B, Morris M, Keijzer R
Analysis of a Parent-Initiated Social Media Campaign for Hirschsprung's Disease
J Med Internet Res 2014;16(12):e288
URL: http://www.jmir.org/2014/12/e288/
doi:10.2196/jmir.3200
PMID:25499427

©Kristy Wittmeier, Cindy Holland, Kendall Hobbs-Murison, Elizabeth Crawford, Chad Beauchamp, Brodie Milne, Melanie Morris, Richard Keijzer. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 11.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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

Therapeutic Affordances of Social Media: Emergent Themes From a Global Online Survey of People With Chronic Pain

Mark Merolli[1], B.Physio (Hons); Kathleen Gray[1], PhD; Fernando Martin-Sanchez[1], PhD

Health and Biomedical Informatics Centre, The University of Melbourne, Melbourne, Australia

Corresponding Author:
Mark Merolli, B.Physio (Hons)
Health and Biomedical Informatics Centre
The University of Melbourne
Level 1
202 Berkeley St
Melbourne, 3010
Australia
Phone: 61 390354943
Fax: 61 390358873
Email: mmerolli@student.unimelb.edu.au

Abstract

Background: Research continues to present tenuous suggestions that social media is well suited to enhance management of chronic disease and improve health outcomes. Various studies have presented qualitative reports of health outcomes from social media use and have examined discourse and communication themes occurring through different social media. However, there is an absence of published studies examining and unpacking the underlying therapeutic mechanisms driving social media's effects.

Objective: This paper presents a qualitative analysis thoroughly describing what social media therapeutically affords people living with chronic pain who are self-managing their condition. From this therapeutic affordance perspective, we aim to formulate a preliminary conceptual model aimed at better understanding "how" social media can influence patient outcomes.

Methods: In total, 218 people with chronic pain (PWCP) completed an online survey, investigating patient-reported outcomes (PROs) from social media use. Supplementary to quantitative data collected, participants were also given the opportunity to provide further open commentary regarding their use of social media as part of chronic pain management; 68/218 unique users (31.2%) chose to provide these free-text responses. Through thematic content analysis, 117 free-text responses regarding 10 types of social media were coded. Quotes were extracted and tabulated based on therapeutic affordances that we had previously identified. Inductive analysis was then performed to code defining language and emergent themes central to describing each affordance. Three investigators examined the responses, developed the coding scheme, and applied the coding to the data.

Results: We extracted 155 quotes from 117 free-text responses. The largest source of quotes came from social network site users (78/155, 50.3%). Analysis of component language used to describe the aforementioned affordances and emergent themes resulted in a final revision and renaming of therapeutic affordances: "exploration" (52/155, 33.5% of quotes), "connection" (50/155, 32.3% of quotes), "narration" (33/155, 21.3% of quotes), "adaptation" (13/155, 8.4% of quotes), and "self-presentation" (7/155, 4.5% of quotes). Of the most described affordances, "exploration" was based on a propensity for participants to explain their social media use for information seeking purposes. "Connection" placed greater emphasis on interaction, highlighting themes of "exchanging information" and "mitigating isolation". Responses regarding "narration" highlighted the value of shared experiences and the emotionally cathartic role this plays.

Conclusions: Much of the efficacy of social media may be explicable via a closer examination of therapeutic affordances. Particular areas that warrant attention include social media's ability to filter and guide people to useful information, connect individuals, and share experiences. Further research into a variety of chronic conditions is warranted. Coupled with the results of the present study, a greater theoretical basis detailing how social media may foster health outcomes may lead to an improved evidence base for conducting research and may inform recommendations for social media use in chronic disease management.

(J Med Internet Res 2014;16(12):e284) doi:10.2196/jmir.3494

KEYWORDS

social media; chronic disease; chronic pain; therapeutic affordances; thematic content analysis; patient-reported outcomes; model

XSL•FO
RenderX

Introduction

Background

Studies continue into the therapeutic effects of social media use in chronic disease management, suggesting that social media may be beneficial to people living with chronic disease [1-3]. However, we have not yet gained a detailed knowledge of these effects. Common research approaches used in qualitative studies in the social media domain include, but are not exclusive to, examining the data mined from social network sites (SNS) and online support groups [4,5], thematic analysis of blog content [6], and phenomenological coding of video narratives on YouTube [7,8]. Results commonly describe themes of support, information provision, and online privacy concerns. However, there is relatively little published literature that contextualizes such themes within a model for understanding what underlying mechanisms drive observed effects and health outcomes from social media use.

Previous studies in psychology have examined the relationship between individuals and their environment in an attempt to clarify how interaction may explain the ensuing behaviors of individuals [9]. The theory of "affordances" describes this relationship, which we discuss further in the following section. When interaction with social media mediates health outcomes, the affordance concept can be adapted to examine the underpinning properties of various social platforms affording therapeutic effects. Henceforth, our study refers to "therapeutic affordances" to present an in-depth analysis of such factors. Reference to therapeutic affordances within a health context has previously been seen in only a few studies. Examples include neurological dysfunction and mental health [10-12]. However, the term was contextualized differently. Fetters and Ellis proposed that therapeutic affordances guide action as they "give insight into the relationship between the patient's altered person constraints and the successful accomplishment of meaningful goals" (p. 145) [10].

The Theory of Affordances

Affordance was originally coined by psychologist James Gibson in relation to his work in visual perception. He used the term to describe environmental properties relative to individuals. More technically, affordances are actionable possibilities that exist within the relationship between the actor and features of the environment [13,14]. Donald Norman takes this to a more concrete level, bringing affordance theory into mainstream culture, through its application to the design of everyday things [15]. He stresses the importance of both the actionable possibility and the way that this is conveyed or made visible to the actor [13]. This refined model of affordance leans heavily on the user's experiences, understanding, goals, and past experiences, rather than purely the physical qualities of the environment or object as suggested by Gibson [15]. Hence, Norman suggests that when designing something, every attempt should be made to convey information that intimately outlines the affordance and highlights trouble-free use [15].

Therapeutic Affordances of Social Media

Affordances have been discussed in various academic articles in different contexts [13,16,17]. However, the way they are described and related back to the work of Gibson and Norman differs considerably. Relative to social media, appropriations of affordances have more recently evolved from human-computer interaction (HCI) research. HCI researchers began using the term "affordances" when studying computer technologies as an artifact. In reference to the present study, social media becomes the artifact. Affordances have included physical affordances (supporting or facilitating physical activity), cognitive (supporting thinking or learning models), sensory (supporting or facilitating sensing), functional (supporting physical activity for a functional purpose), and motivational (supporting motivational needs) [13]. To describe the likely end uses and effects of social media, the focus is placed on the platform, its interface, and the formally described services it provides. Much of the utility of describing affordances of social media in HCI research relates to usability factors because fundamentally, affordances of technology are emergent properties of the integration of the user and the medium [18,19]. However, definitions of social media refer to them as tools designed to allow for the realization of the dynamic interactive qualities of Web 2.0, namely openness, participation, interaction, and collaboration [1,20-22]. Therefore, a simple examination of usability via interaction between user, hardware, and platform interface is not sufficient [13]. In the case of social media as the artifact, it is the user who is ultimately given importance because it is their preferences, attitudes, beliefs, intentions, motivations, experiences, and needs that all converge to allow therapeutic affordances to be realized [13].

Relative to health outcomes, thought needs to go into how social media are being used and in what context. Therefore, one of the preliminary tasks is to identify the type of user being targeted and the environment in which social media are to be used. This provides insight into the user's motivations for using social media. In the present study, the context is people with chronic pain (PWCP) self-managing their condition. Referring to Norman's approach to affordances is thus helpful in this context, as it suggests that different users will approach, understand, and react to the same object differently [15].

Aim

Building on the foundation of an extensive literature review [23], the aim of this paper is to understand how a more targeted approach to social media use by considering therapeutic affordances may better address the needs of people living with and self-managing chronic disease. We theorize that different social media interactions can precipitate different health effects for different people. We anticipate that this may guide clinicians' decision making surrounding whether social media may form a meaningful part of patient management and also may guide further research in this field.

XSL•FO
RenderX

Methods

Overview

This paper reports on findings from the analysis of qualitative data collected as part of a global online survey of PWCP, investigating patient-reported outcomes (PROs) from their social media use to manage chronic pain (Figure 1). This follows approaches to developing and testing conceptual models in health [24]. The Human Research Ethics Committee at the University of Melbourne approved this study (ID No. 1339414).

Figure 1. Screenshot from chronic pain online survey recruitment video (prepared with Animoto).



Rationale for Qualitative Study

Previous research on social-mediating technologies in health has employed thematic content analysis (TCA) to examine qualitative data [6-8,25-29]. Data are collected from a given set of participants reflecting on their experiences within a particular area of study to identify themes [30,31]. The researcher's epistemological stance is an objective one, where themes are extrapolated from the data in an attempt to give meaning to the commonality of collective participants [30]. These data can be coded "inductively", meaning themes are pulled from the data gradually, or "deductively", whereby themes are imposed on the data at the outset as a way of approaching it [31]. When conducting inductive coding, grounded theory is used to describe the process whereby categories are identified as they emerge from the data [31]. These themes are given names or labels in a manner that reflects the words of the participants as a whole. In reality, in the realm of phenomenological research, thematic content analysis requires the researcher to identify meaning within themes and then further put this into the context of the study. The TCA process is not a complete analysis until this has occurred. While open to more interpretative bias, TCA adds another dimension to health research, particularly research with substantial social components. Hence, it is a good fit for the present study. Unlike quantitative methods, which can still be applied to qualitative data, TCA allows the findings to be viewed from the participant's perspective, in their own words. This is particularly useful when investigating a complex construct, such as social media [31].

Recruitment and Data Collection

Adults (18 years or older), with chronic pain (3 months or greater), who used social media as part of their self-management were invited to participate in this online survey via online social networks. Searching Google periodically from March 1 through to May 20, 2013, terms such as "online health networks", "online pain support communities", "chronic pain organizations", "chronic disease organizations", and "international pain organizations" were used to identify potential recruitment channels. Searching was limited to English language. Also included were common social networks, and we targeted active chronic pain groups, such as those on Facebook, Twitter, Daily Strength, and PatientsLikeMe. We also contacted various other influencers (at the support group/organizational level and individual level based on word of mouth).

Requests for assistance were emailed to each organization/group's moderator. We made it clear that we were focusing the survey on "pain interference" as a result of living with chronic disease. The email contained a link directly to the survey, with the plain language statement and informed consent appearing on the survey webpage. This allowed moderators to review the suitability of the research to their members. A recruitment video was also created by the study's primary investigator to complement the email, and a link to the video was pasted into the email text [32]. If the moderator was willing to post the survey to their members, a link to the survey was placed on the websites of the groups, shared on social media, and included in newsletters where appropriate. This emerging approach to recruitment using social media is documented in

XSL•FO
RenderX

the literature [33-36] and was the subject of an article published elsewhere [37]. Given that we relied on a viral dissemination of the survey link via social media and did not invite participants directly via individual emails, it was impossible to calculate response rate based on number of invitations compared to responses.

The survey was conducted using the online survey software, SurveyMonkey. The survey was open from May 21 to June 30, 2013. Participants were asked to provide quantitative and qualitative data in a variety of areas: demographics, health/pain status, social media used, therapeutic affordances, and PROs from use. The full survey instrument can be found in Multimedia Appendix 1.

The focus here is on free-text responses from this survey. Supplementary to quantitative data collected (which is the focus of a separate paper), participants were given opportunities at several points within the survey to provide open commentary relating to the social media platforms they used as part of chronic pain management. The social media they could comment on included SNS, blogs, wikis, microblogs, virtual worlds, tagging/aggregation sites, video sharing sites, photo sharing sites, chat rooms, and discussion forums. An example of how the open commentary questions were phrased was “Please use this space if you would like to comment further about your use of BLOGS for your chronic pain self-management”.

Participants

In total, 231 PWCP consented to take part in the survey; 4 of these supplied no further information and a further 9 answered “no” to the question, “Do you have chronic pain?” The final dataset thus represented 218 completed surveys. A subset of all participants (68/218, 31.2%) chose to provide free-text responses. Table 1 shows their demographic characteristics. The data highlight a majority of married/partnered females, not working because of ill health. However, a wide range of ages and education levels are represented.

Table 1. Participant demographics (n=68).

Characteristics	n (%)
Gender	
Male	10 (15)
Female	58 (85)
Age range	
18-29	14 (21)
30-39	14 (21)
40-49	16 (24)
50-59	16 (24)
60+	8 (12)
Marital status	
Never married	17 (25)
Currently married/partnered	41 (60)
Separated/divorced/widowed	10 (15)
Level of education	
High school or less	25 (37)
College/University completed	26 (38)
Post-graduate degree completed	17 (25)
If not working for pay (reason?) (n=40)	
Ill health	34 (85)
Other	6 (15)

Data Analysis

Dataset

The dataset for analysis was a total of 117 free-text responses given by these 68 participants; 155 separate quotes were extracted. Each quote remained linked to the respondent with a unique identifier. The quotes pertained to 10 different types of social media. SNS accounted for the majority of quotes (78/155, 50.3%), followed by much smaller samples. The full list can be seen in Table 2.

XSL•FO
RenderX

Table 2. Dataset based on platform.

Platform	Used each platform (N=218), n	Free-text responses collected, n	Supplying free-text responses, n (%)	Extracted quotes, n	Total quotes (N=155), n (%)
Social network sites	189	58	58/189 (31)	78	78 (50.3)
Discussion forums	86	18	18/86 (21)	16	16 (10.3)
Blogs	88	11	11/88 (13)	14	14 (9.0)
Wikis	74	10	10/74 (14)	10	10 (6.5)
Videosharing sites	60	7	7/60 (12)	10	10 (6.5)
Photosharing sites	18	5	5/18 (28)	9	9 (5.8)
Microblogs	29	6	6/29 (21)	6	6 (3.9)
Chat rooms	11	3	3/11 (27)	5	5 (3.2)
Virtual worlds	7	2	2/7 (29)	4	4 (2.6)
Tag/Aggregators	12	2	2/12 (17)	1	1 (0.6)

Procedure

Next, 3 investigators used TCA to examine each quote, developed the coding scheme, and applied the coding to the data. Initially, quotes were coded deductively according to their best fit to a set of therapeutic affordances of social media that we had previously identified: identity, flexibility, structure, narration, and adaptation [23]. Then, an alternative inductive analysis was performed to group quotes into other possible themes more closely based on the language components of the quote. Following this two-step coding process, we reviewed and renamed our previous set of therapeutic affordances. A step-by-step workflow can be seen in Figure 2.

In Phase 1, using the “analyze data” function in SurveyMonkey, the free-text responses for each platform were accessed. The first step was to calculate the number of responses provided. The next step in this analysis was to read each response individually and create preliminary categories in SurveyMonkey based on the therapeutic affordances: identity, flexibility, structure, narration, and adaption. Any responses containing quotes referring to two or more therapeutic affordances were categorized as “mixed”. The coding procedure was challenged by responses that appeared to be between categories and those containing descriptions that were outside the proposed categories.

In Phase 2, based on the affordances applied in Phase 1, responses were syphoned to extract meaningful quotations that pertained to the respective therapeutic affordance. In several instances, responses contained more than one quote spanning more than one affordance. The quotes were then tabulated to create a better visualization of the data (Multimedia Appendix 2). A picture emerged of the prevalence of each affordance described by participants. Quotes were tabulated according to therapeutic affordance, social media type, and whether the connotation was positive or negative.

In Phase 3, all quotes were revisited several times to ensure consistency of categorization. This process was repeated until no further re-categorization was necessary.

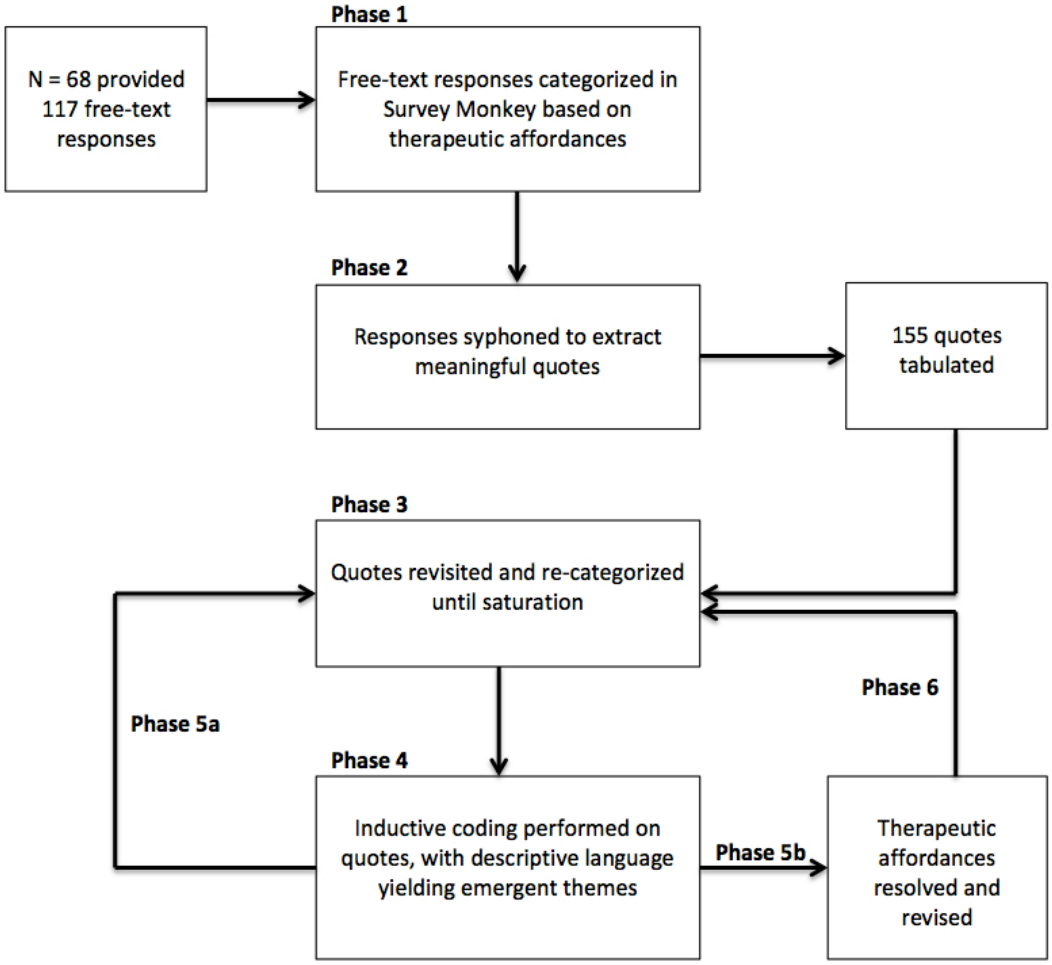
In Phase 4, the coding method was altered, focusing on inductively analyzing the data. Descriptive language in each quote was used to help formulate themes emergent from within the data that captured the essence of what the quote described. This process was performed independently by all 3 researchers and repeated until no new themes emerged. Themes were discussed and compared for overlap. Any consistent themes were automatically coded. This process yielded 15 themes that captured the essence of each therapeutic affordance.

In Phase 5, the entire dataset was revisited (5a). The descriptive language and themes coded from the quotes describing each affordance in Phase 4 were reconsidered and grouped, thus resolving into a revised set of 5 therapeutic affordances: self-presentation, connection, exploration, narration, and adaptation (5b).

The process in Phase 5 also helped to code those quotes that did not fit neatly within one category or those that initially appeared to span multiple categories. Thus, Phase 3 was once again revisited until Phase 6, which revalidated the newly coined affordances, was complete.

XSL•FO
RenderX

Figure 2. Study workflow.



Results

Therapeutic Affordances Described

Overview

Based on the 155 extracted quotes, responses were mostly positive (134/155, 86.5%) (Multimedia Appendix 2). The five finalized therapeutic affordances appeared in the data to varying degrees. The largest number of quotes related to the "exploration" (52/155, 33.5%) and "connection" (50/155, 32.3%) affordances, followed by "narration" (33/155, 21.3%), then "adaptation" (13/155, 8.4%), and "self-presentation" (7/155, 4.5%). Key language used by participants to describe each affordance and the inductively identified themes are presented and discussed in Table 3. The revised therapeutic affordances are then presented. The most representative quotes were extracted and examples highlighted to demonstrate the themes in question.

XSL•FO
RenderX

Table 3. Descriptive language, inductive themes, and validated therapeutic affordances.

Language	Theme	Therapeutic affordance
Disclosure	Disclosure	Self-presentation
Control	Control	
Private	Identification	
Anonymity		
Masquerade		
Communicate	Interaction	Connection
Exchange	Exchanging information	
Share	Support	
Discuss	Mitigating isolation	
Advice	Geographic freedom	
Support		
Isolation		
Alone		
Worldwide		
Look	Information seeking	Exploration
Research	Learning	
Find	Reputability	
Learn		
Monitor		
Moderate		
Share	Imparting knowledge	Narration
Stories	Understanding	
Experiences	Emotional catharsis	
Journey		
Distress		
Flared-up	Variation in use	Adaptation
Affected		
Bad		
Sick		
Hospital		

Self-Presentation: Disclosure, Control, and Identification

This affordance was present in the responses participants gave regarding the level of information they presented to the world via social media. It also highlighted user preferences surrounding identifiable versus anonymous interactions. Participants did not appear to place a great emphasis on the differing self-presentations afforded by social media. However, the quotes obtained showed a relatively equal share of positive and negative perceptions. Information disclosure was a topical theme discussed in terms of self-presentation. Participants showed a propensity to value private interaction and to limit the exposure of their online activity: “The discussion forum we use is a private support group, so only other sufferers of my condition are able to participate” [resp. 204].

This was reiterated more strongly in one person’s perception of potentially negative consequences from disclosing personal information. Powerful language included words like “fear” and “masquerading”. This same participant described the apprehension attached to personal disclosure. They felt strongly about their personal information being used against them: “My fears around disclosure and how it impacts psychologically physically and emotionally is enough to prevent me from disclosing in the social media” [resp. 64] and “I do not want to disclose my personal and painful journey via a social network site for it to be highlighted by others and ‘used’ as a way to finish me in my job” [resp. 64].

However, positive opinion describing “self-presentation” contained a sense of control and autonomy found in social media use: “This is an excellent medium for me to be able to control

XSL•FO
RenderX

my social interactions" [resp. 120], "Social network sites allow me to interact with others on my own terms" [resp. 184], and "writing a decent-ish blog can allow you to be the person you were to an extent" [resp. 229].

Participants had differing preferences towards whether others knew their identity. In one instance, one participant described a desire to remain anonymous to friends and family, yet known to other PWCP: "I find these boards are more intimate if I can remain anonymous to family and friends but known to others on the board" [resp. 176].

Conversely, anonymity was also used to highlight the potential hazards surrounding self-presentation and an inability to verify identity through social media: "I see individuals masquerading as members of the medical profession" [resp. 191].

Connection: Interaction, Exchanging Information, Support, Mitigating Isolation, and Geographic Freedom

An ability to connect with others featured frequently in the responses. Participants commonly used social media to reach out to others in similar situations, share/exchange information, and offer support. A key objective was mitigating isolation. This was aided by a sense of freedom afforded by social media to interact when and where it suits.

Connection to others was one activity particularly suited to social media. Participants commonly described their interactions enabled by social media: "We are able to help each other with practical advice about drug side effects etc" [resp. 83], "Having immediate access to information and people is enlightening" [resp. 206], and "Talking to people who understand is great" [resp. 180].

Participants reported actively sharing and exchanging information. The words "sharing" and "exchanging" of information appeared frequently throughout the responses, with words like "valuable" often used in conjunction. Sharing of information ranged from general disease-specific information and advice, through to practical tips. On a more personal level, participants tied the ability to exchange information to an enhanced feeling of self-worth and validation, often inhibited by living with chronic illness: "I feel valued sharing my resources with others" [resp. 79], "I focus on sharing what knowledge or information I have in other areas, this mitigates the constant feeling of failure that comes along with ill health and the inability to be more active" [resp. 177], and "The exchange of information available is invaluable" [resp. 38].

Like all therapeutic affordances, connection seen via information exchange is not immune to criticism. One person who pondered the reputability of information exchanges expressed this as "Unmoderated sites are potentially dangerous, unqualified ppl routinely diagnose, recommend treatment, use research findings inappropriately, misinform other ppl who may not know any better" [resp. 8].

Providing and receiving support was described as another valuable utility of social media. A sense of camaraderie enabled by social media was outlined, bringing together people in similar situations. Support fostered through interactions was emotionally cathartic to participants and gave them the motivation to persist with management: "only a person with similar problems really gets what you're going through" [resp. 201], "we all support each other during our tough times and still communicate even during our good days" [resp. 140], and "Being able to communicate with people can be up lifting and cathartic" [resp. 229].

The strength of social media to connect was evident as analysis uncovered strong positive sentiments towards social media use as part of chronic pain management: "I find general support from others with the same condition on these sites has stopped my suicide attempts" [resp. 217], "[social network sites] actually saved my life. The people I connected with were other with MS…we all support each other" [resp. 140], and "The friendships and information I've gained from those have been life saving—probably in the literal sense" [resp. 71].

On the other hand (and although only a small representation), those who perceived "connection" negatively discussed a sense of withdrawal and frustration: "Social networking can sometimes have a negative impact ie withdrawal from real life" [resp. 56] and "Frustrating and at times more painful trying to explain to strangers" [resp. 124].

Mitigating isolation was the most present theme describing "connection", accounting for 28% (14/50) of quotes. Participants felt extremely positive towards social media's ability to keep them from being dominated by loneliness. Feeling connected and knowing they were not alone were apparent effects. The words "alone" and "isolation" appeared six and eight times respectively: "Without social networking I would feel far more alone when my conditions are bad" [resp. 46], "pain is isolating in real life and social networking can help to reduce the amount of isolation" [resp. 56], and "Social network sites have allowed me to have a social life…when the pain is bad, which is frequent, I cannot leave my house and spend time with friends" [resp. 196].

A final element underlying "connection" was seen through the abundant quotes detailing reach and spread of communication afforded by social media. Participants freely commented that their use of social media provided borderless interactions, thus highlighting that managing chronic disease can span the world. Social media allows these people the opportunity to extend their networks: "[SNS] is often putting patients in touch with treating practitioners around the world" [resp. 79], "people all over the world can communicate and support others in their same situation" [resp. 78], and "I live in a regional area and helpful Dr. services are not available! I'm discovering the positives of using the sites" [resp. 164].

Geographic freedom has also increased the availability of support. Being able to connect with fellow PWCP internationally means that time of day need no longer be a hurdle: "now I have friends worldwide…they are there day or night if I need" [resp. 71] and "are mostly helpful due to the 'you are not alone' value" [resp. 215].

Exploration: Information Seeking, Learning, and Reputability

"Exploration" was the most noted therapeutic affordance based on number of quotes. Survey participants frequently described

XSL•FO
RenderX

searching for information. A variety of search-related activities were reported

Above all, participants described using social media for guidance towards useful information for pain management. At times the search for information yielded resources that filtered further useful resources. Participants reported using social media to seek out information on a wide variety of matters relevant to their condition. This included disease-specific information, research, and treatments: "I selectively filter through information...I look for information around the management of pain and adapting in certain situations" [resp. 51], "The blogs of health researchers and professionals are very important tools for finding new information, current research etc" [resp. 217], and "Practical tips, medication, treatment recommendations have often lead me to helpful sites" [resp. 190].

Some participants valued the information found through social media: "Gives me more information about a specific condition or more things to discuss with my doctor" [resp. 200]. Again, some voiced cautions. Using the example of wikis, two participants highlighted: "Wikis are not always particularly useful, as entries tend to be very general and not always up to date" [resp. 190] and "Seem to be most prone to bias…so bad information doesn't get corrected" [resp. 125].

Another process was that of "learning" resulting from use. Participants declared that search and guidance promoted by social media were useful to aid learning about managing pain, disease-specific knowledge, and self-management: "Have particularly used social networking to learn tips about managing/avoiding pain" [resp. 112], "…use blogs so as I can learn as much as I can about my condition. Knowledge is power, etc" [resp. 227], and "They have definitely helped my understanding of nerve pain and pain in general. I have grasped anatomical and medical info much easier with YouTube" [resp. 229].

From a more cautious perspective, quality of information resonated with some survey participants: "Value depends on quality of information" [resp. 190]. Participants consider this pivotal to the success of self- management via social media. One participant described the free flow of information made available by social media: "The beauty of social media is that there is no censor or control so the right information gets through" [resp. 78].

However, this conflicts somewhat with the perceptions that information made available via social media can at times be disreputable: "…can be misleading depending on who has written the article" [resp. 66] and "Seem to be most prone to bias...so bad information doesn't get corrected" [resp. 125].

This may be addressed via careful consideration of individual preferences for moderated and monitored interactions, as seen in many quotes: "I much prefer a medically knowledgeable, balanced platform such as a monitored [one] where good, sensible advice and help is offered" [resp. 40] and "one becomes discriminating with the types of boards and if the moderator is a good one, this sets the tone of the board and generally adds to the value of the information" [resp. 190].

Narration: Imparting Knowledge, Understanding, and Emotional Catharsis

The narrative affordance was well documented in the results of this survey. Multiple instances of the positive effect of sharing experiences via social media were described. This was true of both sharing one's own experiences and accessing another's experiences. Accessing other's experiences was most notable, accounting for 48.5% (16/33) of total "narration" quotes.

Where more active participation was described, imparting knowledge by sharing one's own experiences with chronic pain fostered a sense of personal satisfaction and self-worth. This made participants feel useful, as they were able to impart knowledge they had accumulated: "networking sites make me feel useful as I can share my stories with newly diagnosed people" [resp. 185], "I write reflective and creative pieces about my journey in general...often I won't even discuss the physical side of it, but rather the lessons learned" [resp. 189], and "It is very helpful to share stories, symptoms, and problems with other sufferers" [resp. 223].

Participants also described that accessing the experiences of others played a beneficial role in improving understanding. Through accessing experiences and narratives about pain from other people, participants felt able to better manage their condition: "I find it helpful to know how other people manage their own pain" [resp. 191] and "having people who can make comments and talk about their stories give me more information about what other people are dealing with" [resp. 32].

However, narration was most commonly discussed as a means to manage the emotional burden of living with chronic pain. Participants discussed the catharsis, comfort, and validation from accessing other people's experiences: "It just helps to know there are other people in the same boat who work hard at their disorder and succeed. It gives me hope and patience to keep trying" [resp. 212], "Seeing others who are like me, is a HUGE comfort" [resp. 190], and "Many of my own experiences were also shared by others. Often it is quite validating to see others have been thru the same things I have" [resp. 190].

Conversely, some did caution that the narrative experience could be counterproductive, as people may be potentially distressed from sharing experiences: "I do not feel able to share my experience as yet as I become very distressed when I think/write/talk about my condition" [resp. 47], "Sometimes reading other peoples journeys is useful but I do not like to immerse myself in too many down stories as it makes me worse so I am careful about what I read" [resp. 209], and "I find most blogs very distressing as I 'see' how my future may progress" [resp. 51].

One participant took this further and described the potential to become alienated by sharing experiences online: "Unfortunately it can often feel that when people read your posts in regard to chronic pain, it can create speculation about the condition...it can make me feel, especially being a young male, that I am just a whinger and need to toughen up" [resp. 84].

In a similar light, narratives may also provide greater scope to malinger and catastrophize one's condition: "I do find many social network chronic pain sites regular participants too needy.

XSL•FO
RenderX

I can't abide the constant call for, supply of prayers and the desperation for others to feel/ understand/ get their pain" [resp. 40] and "sites become a platform for particular ppl to use to verbalize complaints and negativity, without any constructive discourse" [resp. 8].

Adaptation: Variation in Use

The "adaptation" affordance captures the way social media enables participants to adapt their self-management behaviors in relation to their condition status and/or needs at particular points in time in various ways. Most apparent was social media's utility in allowing participants to alter their self-management when pain flared or health deteriorated: "I am severely affected at present. I cannot speak, or tolerate the stimulation of people being around me...I am bedridden and need to be in a very low stimulus environment…but youtube is ok for me with care" [resp. 190], "It's nice to be able to learn even when I can't read or sit up" [resp. 27], and "I will post a photo on Instagram if I am in hospital and people want to know how I am going but I am too sick to tell them" [resp. 189].

Discussion

Principal Findings

Refining Therapeutic Affordances

In our previous work, we postulated five therapeutic affordances: "Identity" was used to present perceptions regarding disclosure of identity in online social environments. "Flexibility" described the time-space freedom enabled by social media, such as the ability to interact at a time suiting the individual and wherever one chooses. "Structure" described the guidance and filtration present in information seeking that social media can provide. "Narration" encompassed social media's utility to provide a platform to share stories of illness, and finally, "adaptation" referred to the capacity for one's self-management to evolve through social media to meet particular needs based on current disease status [23].

While our initial nomination of these affordances provided a useful frame for early theorizing, analysis of the qualitative survey data enabled a much richer conceptualization and categorization of the underlying themes describing each therapeutic affordance. Bearing in mind that the use of the term affordance has previously attempted to clarify behaviors an individual may perform, referring to affordances as all actionable possibilities latent in the environment [9], we defined two tests of our affordances in relation to the data from the present study: Does the therapeutic affordance account for the data? Does the therapeutic affordance describe an actionable possibility? Consequently, the data highlighted a lack of clarity in some of our early descriptions. "Narration" and "adaptation" passed the tests; however, the other three affordances were found to be incongruent.

First, the label "identity" is not an actionable term, nor does it describe elements of disclosure or privacy central in participant responses. Quotes described preferences for anonymity, control of one's self-image, the precautions required when interacting online and not being able to verify identity. Therefore, in regards to all of these elements relevant to impression management, "self-presentation" is more descriptive and appropriate than "identity" [38].

Social media are often praised for the "flexibility" they offer health consumers to interact where and when they choose. Responses in the present study also displayed a propensity to praise the opportunity to connect and interact with others (including borderless communication). However, participants also valued supporting one another, mitigating isolation, and exchanging information. Descriptive language was slanted towards communication, discussion, and exchange rather than the flexibility afforded by social media. Thus, the therapeutic affordance is better described as the "connection" of individuals.

"Structure" presented the same semantic challenge. While one can use "structure" as a verb, it is ambiguous. Previous notions that social media guide and filter information seeking still persist. However, the overarching theme from the survey data was "information seeking". Participants used language such as "finding", "looking", and "searching" for information. Hence, "exploration" is more appropriate.

Interaction and sharing seen in "connection" and "narration" embody the very participatory and collaborative nature on which Web 1.0 has evolved into Web 2.0 [1,39]. However, despite "exploration" being the most discussed afforded use of social media (33.5% of total responses), it is important to recognize the language used to describe it (ie, seeking, finding, searching, looking). The language used is information-focused, not user-centered [40,41]. This sentiment has been previously reported [3], indicating that perhaps social media are still not being realized to their full potential as many-to-many communication tools.

Similar difficulties surrounding labeling affordances of social media are apparent in other academic literature. In a study of social media affordances within organizational processes [42], four affordances are based on the authors' perceptions of the activities performed in business processes salient to social media's utilities: visibility, editability, persistence, and association. We suggest that ideas such as "visible" and "persistent", while informative, are contentious in terms of whether they are indeed actionable by people or rather, characteristic of content. These contentions warrant attention by all researchers working with the concept of social media affordances.

Proposing a Theoretical Model

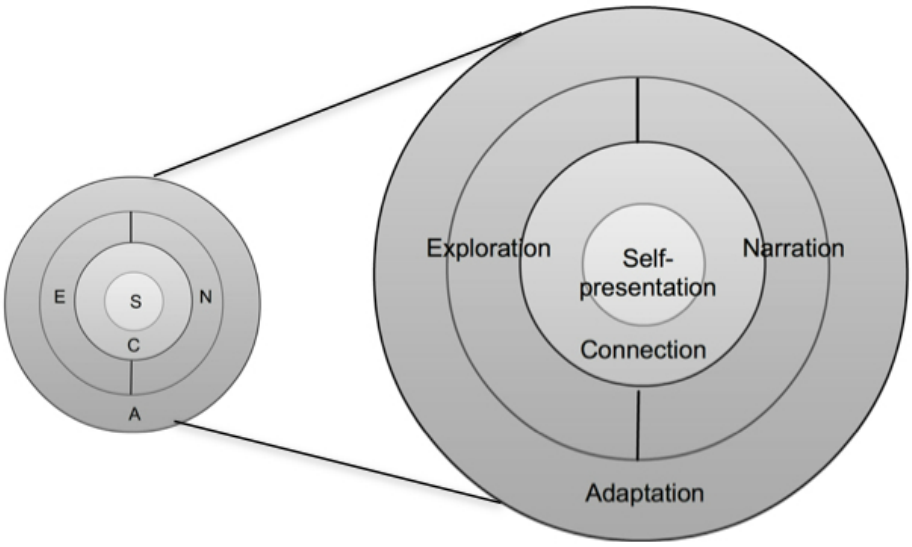
Resulting from this reflection on the analysis presented in this paper, the therapeutic affordances of social media were refined and resolved into self-presentation, connection, exploration, narration and adaptation (SCENA). This resolution enabled us to develop a conceptualization of these affordances relative to health outcomes, not only revising their definitions but also relating them to each other within a new framework that can be visualized as a theoretical model (Figure 3).

The SCENA Model of Therapeutic Affordances of Social Media has at its core preferences and perceptions regarding one's online image or digital identity. "Self-presentation" feeds into the ability of social media to "connect" individuals. The next layer

XSL•FO
RenderX

is shared by “exploration” and “narration”, both of which take into consideration varying preferences for self-presentation and how individuals connect. The outermost layer of the model is “adaptation”, that is, social media use allowing self-management behaviors to adapt to individual needs at given points in time. This will influence and be influenced by the other therapeutic affordances to varying degrees.

Figure 3. SCENA model of therapeutic affordances of social media.



Theorizing About What Makes Therapeutic Affordances Effective

Different categorizations have been proposed as ways to organize affordances of social media. If therapeutic affordances are to be meaningful in a chronic disease management or broader health context, then categorizing them to show how they function to produce effects is a critical contribution that health informatics can make.

The use of classification systems appears in published literature in fields other than health. For example, Day and Lloyd describe technology affordances in an educational context [14], Jordan presents a classification scheme based on psychological theory (physical, psychological, social, cognitive) [43], and Zhang explores information-communication technology (ICT) from a motivation perspective [44]. Other attempts within an ICT context have paired user with design, describing affordances as action, functional, and structural [45].

A classification system proposed in [13] may be more relevant to the present study. The resulting schema is that of physical, cognitive, affective, and control affordances. “Physical” are broadly defined as the physical attributes of social media that the user can manipulate (ie, sound and display settings). “Cognitive” are defined as intellectual attributes aiding or facilitating mental processing, or problem solving. They describe the ability to provide information, reinforcement, and suggestions. “Affective” comes from a psychosocial standpoint, related to the attributes of social media that stimulate emotional responses. Finally, “control” exists in social media’s ability to allow the user to make autonomous choices regarding their interactions. For example, one’s level of self-presentation to others will to some extent be dictated by the degree of control allowed by the platform. Control is described as a powerful affordance funneling positive behavioral change, as users who feel they have more control over how they use social media will generally behave more positively [13].

While the present study has appropriated affordance to describe how social media mediate therapeutic effects, the above categories (in particular, cognitive, affective, and control) may still prove useful as a way to theorize the precise mechanisms through which therapeutic affordances work (Table 4). It is also prudent to note that therapeutic affordances mean different things to different people and interpretation may vary. For this reason, those outlined in the present study are a guide and may serve as a base for further investigation.

Table 4. Categorization of therapeutic affordances.

	Physical	Cognitive	Affective	Control
Self-presentation				x
Connection			x	x
Exploration		x		x
Narration		x	x	
Adaptation			x	x

XSL•FO
RenderX

"Self-presentation" may be seen as a "control" affordance. Through identification of discourse themes highlighting control, identity, and personal disclosure, as well as language such as anonymity and privacy, it is possible that the self-presentation afforded by social media centers around the ability to "control" online interactions.

Themes identified describing "connection" suggest that much of this affordance comes from the ability to give or receive support and interact with others. Participants spoke of the benefits derived from being able to communicate, discuss, advise, and support each other. This had a beneficial impact on sensations of loneliness and isolation, representing a high socio-emotional component. Thus "affective" may be an appropriate categorization. However, the ability to communicate and generate exchanges with others where and when it suits also describes an element of "control" afforded to users, thus "control" is also a suggested classification.

"Exploration" centered much more on the quest to find information, to find, look, and learn, all in the hope of improving disease-specific knowledge. Hence, exploration is possibly best categorized as a "cognitive" affordance. Concurrently, participants had a strong expectation that the information presented on social media be reputable and that in order to do this, moderation and monitoring of online activity is favored. Hence, "exploration" may also fit in the "control" category.

The "narrative" effect speaks primarily of the emotional catharsis that comes with learning from others' experiences and sharing one's own. Hence, suggested categorization as "affective" may be appropriate. However, the strength of the "narration" afforded by social media is also present in how social media allows individuals to develop understanding. This is also true of the personal benefits reported from imparting wisdom to others. For this reason, "narration" may also classify as a "cognitive" affordance.

Finally, the "adaptation" afforded by social media to allow users to evolve their self-management behaviors based on disease-specific needs at different points in time highlights how use changes depending on motivations. For this reason, "affective" is possibly applicable because of the personal or emotional component dictating people's decision-making process. The majority of participants indicated that social media was particularly useful during times of declining health or pain. Social media gave them greater choice of how they interacted with others. Therefore, categorization as a "control" affordance is also suggested.

Limitations

Exploring therapeutic affordances has certain limitations. A 2007 paper addressed difficulties with the affordance concept. In that study, a teacher presented her students with a wiki project [14]. The teacher had learned that a primary affordance of wikis is collaboration, which is commonly described in academic literature about wikis [46]. Despite some students embracing the project, the majority did not. While collaboration is a primary affordance of wikis, this example highlights the challenges in realizing the benefits of affordances under different circumstances. Sometimes other factors present within a given context can interfere with the realization or utility of a given affordance. Modeling and categorizing affordances only go part of the way. The present study does not take into account a plethora of contextual factors that can play a role in determining why therapeutic affordances may explain how social media generate health outcomes. For example, one such prominent factor is the characteristics of the group or individual. Day and Lloyd [14] suggest that these may include previous knowledge, competence, learning preferences, motivation, attitudes, and access to information. When placed in the current context (chronic pain management), the same can be applied [14]. However, particular disease-specific characteristics such as current health status, confounding illness variables, and self-management needs also play a vital part.

Also, the platforms used do not function independently of the context in which they are placed [13]. In regards to this research, it is extremely difficult to infer the most suitable social media to meet particular needs and/or objectives in chronic pain management [13,14,47]. Systematic analysis needs to go beyond an elementary view of the social media available for use and the health outcomes expected. It must consider the ways in which the technology (social media) and factors relevant to the individual interplay and contribute to the ensuing interaction and ultimately, outcomes observed. Only then can the gap between potential for use, actual use, and outcomes be bridged.

Social media presents an extra challenge to health research. Unlike traditional clinical health indicators, it is difficult to measure clinical impact of social media on health outcomes reliably. Validating self-reported health outcomes from social media use still requires effort to establish a sound, reputable evidence base. Formal measurement of PROs is required to assess whether social media use based on therapeutic affordances may be effective for improving health outcomes in chronic disease [48]. In order to do this effectively, future research harnessing information about the therapeutic affordances of social media combined with a validated PROs tool is needed to specifically measure health outcomes [49]. The broader chronic disease landscape in particular pushes us to establish valid PROs measurement research methodologies. This research is currently limited in its transferability to other chronic conditions. We have investigated social media use in relation to chronic pain, with chronic pain being chosen due to it burgeoning effect globally, social stigmatization, and co-existence to and manifestation in a variety of chronic conditions [50]. In order to gain a clearer picture of the utility of therapeutic affordances when examining other chronic conditions (eg, cancer, diabetes, arthritis, depression, fibromyalgia), further research is required. Similarly, it is prudent to acknowledge the influence of self-selection on this research [36,51]. To gain a deeper understanding of the influence of social media's therapeutic affordances, future research may be well served to examine non-responders and perceptions of therapeutic affordances among general social media users in a non-health context [51].

Conclusions

The results of this paper suggest that future social media research in health may be aided by paying closer attention to

XSL•FO
RenderX

therapeutic affordances. By doing so, efficacy may become more discernable. Therapeutic affordances most recognizable in the present study include social media's ability to "explore" information, "connect" people, and "narrate" experiences of illness. Therapeutic affordances are presented in a model to theorize how the interactions enabled by social media may help explain how health outcomes are generated when individuals use social media as part of health self-management. This approach may provide more targeted, evidence-based recommendations for social media use in chronic disease management. Further research to explore for other therapeutic affordances and study across a range of chronic conditions is warranted to build evidence and move one step closer to producing evidence-based guidelines for social media use in chronic disease management.

Acknowledgments

The authors would like to thank the Institute for a Broadband Enabled Society at the University of Melbourne, Victoria, Australia, for ongoing support of this research. We would also like to thank Associate Professor Graham Hepworth of the University's Statistics Consulting Centre for his advice on survey design and analysis and acknowledge advice from Dr Reeva Lederman of the University's Department of Computing and Information Systems about use of the term "affordance".

Conflicts of Interest

None declared.

Multimedia Appendix 1

Survey instrument.

[PDF File (Adobe PDF File), 1MB - jmir_v16i12e284_app1.pdf]

Multimedia Appendix 2

Visualization of full dataset.

[PDF File (Adobe PDF File), 48KB - jmir_v16i12e284_app2.pdf]

References

1. Stellefson M, Chaney B, Barry AE, Chavarria E, Tennant B, Walsh-Childers K, et al. Web 2.0 chronic disease self-management for older adults: a systematic review. J Med Internet Res 2013;15(2):e35 [FREE Full text] [doi: 10.2196/jmir.2439] [Medline: 23410671]
2. Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. J Med Internet Res 2013;15(4):e85 [FREE Full text] [doi: 10.2196/jmir.1933] [Medline: 23615206]
3. Thackeray R, Crookston BT, West JH. Correlates of health-related social media use among adults. J Med Internet Res 2013;15(1):e21 [FREE Full text] [doi: 10.2196/jmir.2297] [Medline: 23367505]
4. Hale TM, Pathipati AS, Zan S, Jethwani K. Representation of health conditions on facebook: content analysis and evaluation of user engagement. J Med Internet Res 2014;16(8):e182 [FREE Full text] [doi: 10.2196/jmir.3275] [Medline: 25092386]
5. Struik LL, Baskerville NB. The role of Facebook in Crush the Crave, a mobile- and social media-based smoking cessation intervention: qualitative framework analysis of posts. J Med Internet Res 2014;16(7):e170 [FREE Full text] [doi: 10.2196/jmir.3189] [Medline: 25016998]
6. Marcus, Westra H, Eastwood J, Barnes K, Mobilizing Minds Research Group. What are young adults saying about mental health? An analysis of Internet blogs. J Med Internet Res 2012;14(1):e17 [FREE Full text] [doi: 10.2196/jmir.1868] [Medline: 22569642]
7. Gómez-Zúñiga B, Fernandez-Luque L, Pousada M. ePatients on YouTube: Analysis of Four Experiences From the Patients' Perspective. Med 2.0 2012;1(1):e1. [doi: 10.2196/med2.2039]
8. Chou WY, Hunt Y, Folkers A, Augustson E. Cancer survivorship in the age of YouTube and social media: a narrative analysis. J Med Internet Res 2011;13(1):e7 [FREE Full text] [doi: 10.2196/jmir.1569] [Medline: 21247864]
9. Gibson J. The ecological approach to visual perception. Hillsdale, NJ: Lawrence Erlbaum Associates; 1986.
10. Fetters L, Ellis T. A perception-action framework for physical therapy for persons with neurologic dysfunction: use of therapeutic affordance and unitless ratio. J Neurol Phys Ther 2006 Sep;30(3):142-147. [Medline: 17029657]
11. Chen YP, Mitch A, Chaffin K, Sargent R. The impact of body-scaled Information on grasping action in toddlers with and without Down Syndrome. The Internet Journal of Allied Health Sciences and Practice 2013;11(3):1540-1580.
12. Duff C. Exploring the role of 'enabling places' in promoting recovery from mental illness: a qualitative test of a relational model. Health Place 2012 Nov;18(6):1388-1395. [doi: 10.1016/j.healthplace.2012.07.003] [Medline: 22853876]

XSL•FO
RenderX

13. Zhao Y, Liu J, Tang J, Zhu Q. Conceptualizing perceived affordances in social media interaction design. Aslib Proceedings: New Information Perspectives 2013;65(3):0001-0253.
14. Day D, Lloyd MM. Affordances of online technologies: More than the properties of the technology. Australian Educational Computing 2007;22(2):0816-9020.
15. Norman D. The design of everyday things. New York: Basic Books; 2002.
16. Sutcliffe AG, Gonzalez V, Binder J, Nevarez G. Social mediating technologies: Social affordances and functionalities. International Journal of Human-Computer Interaction 2011 Nov;27(11):1037-1065. [doi: 10.1080/10447318.2011.555318]
17. Leonardi PM. When does technology use enable network change in organizations? A comparative study of feature use and shared affordances. MIS Q 2013;37(3):749-776.
18. Bevan N. Extending Quality in Use to Provide a Framework for Usability Measurement, in Human Centered Design. In: Kurosu M, editor. Human Centered Design: First International Conference, HCD 2009. Berlin: Springer; 2009:-.
19. Bradner E. Social affordances of computer-mediated communication technology: understanding adoption. In: Extended Abstracts on Human Factors in Computing Systems. 2001 Presented at: CHI 2001; March 31, 2001; Seattle, Washington p. 67-68. [doi: 10.1145/634067.634111]
20. Van De Belt TH, Engelen LJ, Berben SA, Schoonhoven L. Definition of Health 2.0 and Medicine 2.0: a systematic review. J Med Internet Res 2010;12(2):e18 [FREE Full text] [doi: 10.2196/jmir.1350] [Medline: 20542857]
21. Wietske van O. Organizational Social Media: A Comprehensive Framework and Research Agenda. 2013 Presented at: 46th Hawaii International Conference on System Sciences; 2013; Hawaii. [doi: 10.1109/HICSS.2013.439]
22. O'Reily TB. Web Squared: Web 2.0 Five Years On. 2009. URL: http://www.webcitation.org/6UNsjAlKw [accessed 2014-11-26]
23. 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. [doi: 10.1016/j.jbi.2013.04.010] [Medline: 23702104]
24. Earp JA, Ennett ST. Conceptual models for health education research and practice. Health Educ Res 1991 Jun;6(2):163-171. [Medline: 10148689]
25. Chen AT. Exploring online support spaces: using cluster analysis to examine breast cancer, diabetes and fibromyalgia support groups. Patient Educ Couns 2012 May;87(2):250-257. [doi: 10.1016/j.pec.2011.08.017] [Medline: 21930359]
26. Greene JA, Choudhry NK, Kilabuk E, Shrank WH. Online social networking by patients with diabetes: a qualitative evaluation of communication with Facebook. J Gen Intern Med 2011 Mar;26(3):287-292 [FREE Full text] [doi: 10.1007/s11606-010-1526-3] [Medline: 20945113]
27. Schulz PJ, Rubinelli S, Mariotti G, Keller N. Meeting the ranging of informational needs of chronic low back pain sufferers: conceptual design and rationale of the interactive website ONESELF. Disabil Rehabil 2009;31(25):2118-2124. [doi: 10.3109/09638280902943231] [Medline: 19888842]
28. Rodham K, McCabe C, Blake D. Seeking support: An interpretative phenomenological analysis of an Internet message board for people with Complex Regional Pain Syndrome. Psychol Health 2009 Jul;24(6):619-634. [doi: 10.1080/08870440802563245] [Medline: 20205016]
29. van Uden-Kraan CF, Drossaert CH, Taal E, Shaw BR, Seydel ER, van de Laar MA. Empowering processes and outcomes of participation in online support groups for patients with breast cancer, arthritis, or fibromyalgia. Qual Health Res 2008 Mar;18(3):405-417. [doi: 10.1177/1049732307313429] [Medline: 18235163]
30. Anderson R. Thematic Content Analysis (TCA): Descriptive Presentation of Qualitative Data. Palo Alto, CA: Institute of Transpersonal Psychology; 2007. URL: http://www.wellknowingconsulting.org/publications/pdfs/ThematicContentAnalysis.pdf [accessed 2014-11-27] [WebCite Cache ID 6UPDxjZIC]
31. Pope C, Ziebland S, Mays N. Qualitative research in health care. Analysing qualitative data. BMJ 2000;320(7227):114-116. [doi: 10.1136/bmj.320.7227.114]
32. Merolli M. An analysis of social media use by people with chronic pain. 2013. URL: http://animoto.com/play/EY5Vh6FUHKmeZrqLAIUNQQ [accessed 2014-10-02] [WebCite Cache ID 6T0xAoCxk]
33. Close S, Smaldone A, Fennoy I, Reame N, Grey M. Using information technology and social networking for recruitment of research participants: experience from an exploratory study of pediatric Klinefelter syndrome. J Med Internet Res 2013;15(3):e48 [FREE Full text] [doi: 10.2196/jmir.2286] [Medline: 23512442]
34. Ramo DE, Prochaska JJ. Broad reach and targeted recruitment using Facebook for an online survey of young adult substance use. J Med Internet Res 2012;14(1):e28 [FREE Full text] [doi: 10.2196/jmir.1878] [Medline: 22360969]
35. Yuan P, Bare MG, Johnson MO, Saberi P. Using online social media for recruitment of human immunodeficiency virus-positive participants: a cross-sectional survey. J Med Internet Res 2014;16(5):e117 [FREE Full text] [doi: 10.2196/jmir.3229] [Medline: 24784982]
36. Alshaikh F, Ramzan F, Rawaf S, Majeed A. Social network sites as a mode to collect health data: a systematic review. J Med Internet Res 2014;16(7):e171 [FREE Full text] [doi: 10.2196/jmir.3050] [Medline: 25048247]
37. Merolli M, Martin-Sanchez F, Gray KF. Social Media and Online Survey: Tools for Knowledge Management in Health Research. In: Conferences in Research and Practice in Information Technology (CRPIT). 2014 Presented at: Seventh

XSL•FO
RenderX

Australasian Workshop on Health Informatics and Knowledge Management (HIKM); 2014; Auckland, New Zealand p. 21-29.
38. Orsatti J, Riemer K. Identity and Self-Presentation: from a Representational to a Performative Lens in Studying Social Media Engagement in Organisations. 2012 Presented at: 23rd Australasian Conference on Information Systems; Dec. 3, 2012; Geelong, Victoria p. 12.
39. Kamel Boulos MN, Wheeler S. The emerging Web 2.0 social software: an enabling suite of sociable technologies in healthhealth care education. Health Info Libr J 2007;24(1):2-23. [doi: 10.1111/j.1471]
40. Kaplan AM, Haenlein M. Users of the world, unite! The challenges and opportunities of Social Media. Business Horizons 2010 Jan;53(1):59-68. [doi: 10.1016/j.bushor.2009.09.003]
41. Boyd DM, Ellison NB. Social Network Sites: Definition, History, and Scholarship. Journal of Computer-Mediated Communication 2008;13(1):210-230. [doi: 10.1111/j.1083-6101.2007.00393.x]
42. Treem JW, Leonardi PM. Social Media Use in Organizations: Exploring the Affordances of Visibility, Editability, Persistence, and Association. Communication Yearbook 2012;36:143-189.
43. Jordan P. Designing pleasurable products: an introduction to the new human factors. London: Taylor & Francis; 2000.
44. Zhang P. Motivational affordances: Reasons for ICT design and use. Commun. ACM 2008 Nov 01;51(11):145-147. [doi: 10.1145/1400214.1400244]
45. Grange C, Benbasat I. A Guiding Framework for Developing Theories Investigating the Design Drivers of IT Use and Value. In: Proceedings of the JAIS Theory Development Workshop. 2011 Presented at: ICIS 2011; Dec. 4-7, 2011; Shanghai, China p. 1-28.
46. Mathew D, McKibbon KA, Lokker C, Colquhoun H. Engaging with a Wiki related to knowledge translation: a survey of WhatisKT Wiki users. J Med Internet Res 2014;16(1):e21 [FREE Full text] [doi: 10.2196/jmir.3001] [Medline: 24449712]
47. Merolli M, Gray K, Martin-Sanchez F. Developing a framework to generate evidence of health outcomes from social media use in chronic disease management. Med 2 0 2013;2(2):e3 [FREE Full text] [doi: 10.2196/med20.2717] [Medline: 25075238]
48. Baldwin M, Spong A, Doward L, Gnanasakthy A. Patient-reported outcomes, patient-reported information: from randomized controlled trials to the social web and beyond. Patient 2011;4(1):11-17 [FREE Full text] [doi: 10.2165/11585530-000000000-00000] [Medline: 21766890]
49. Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, IMMPACT. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain 2005 Jan;113(1-2):9-19. [doi: 10.1016/j.pain.2004.09.012] [Medline: 15621359]
50. Access Economics. The high price of pain: the economic impact of persistent pain in Australia. Access Economics 2007:101.
51. Khazaal Y, van Singer M, Chatton A, Achab S, Zullino D, Rothen S, et al. Does self-selection affect samples' representativeness in online surveys? An investigation in online video game research. J Med Internet Res 2014;16(7):e164 [FREE Full text] [doi: 10.2196/jmir.2759] [Medline: 25001007]

Abbreviations

HCI: human-computer interaction
PROs: patient-reported outcomes
PWCP: people with chronic pain
SNS: social network sites
TCA: thematic content analysis

Edited by G Eysenbach; submitted 28.04.14; peer-reviewed by L Fernandez-Luque, J West; comments to author 17.08.14; revised version received 28.09.14; accepted 28.10.14; published 22.12.14.

Please cite as:
Merolli M, Gray K, Martin-Sanchez F
Therapeutic Affordances of Social Media: Emergent Themes From a Global Online Survey of People With Chronic Pain
J Med Internet Res 2014;16(12):e284
URL: http://www.jmir.org/2014/12/e284/
doi:10.2196/jmir.3494
PMID:25533453

©Mark Merolli, Kathleen Gray, Fernando Martin-Sanchez. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 22.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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

eMental Health Experiences and Expectations: A Survey of Youths' Web-Based Resource Preferences in Canada

Felicia M Wetterlin[1]; Marissa Y Mar[2], BA; Erika K Neilson[2], MPH; Gregory R Werker[2,3], MEng, PhD; Michael Krausz[1,2,4], MD, PhD

[1]Institute of Mental Health, Department of Psychiatry, University of British Columbia, Vancouver, BC, Canada

[2]Centre for Health Evaluation and Outcome Sciences, University of British Columbia, Vancouver, BC, Canada

[3]Sauder School of Business, University of British Columbia, Vancouver, BC, Canada

[4]School of Population and Public Health, University of British Columbia, Vancouver, BC, Canada

Corresponding Author:
Felicia M Wetterlin
Institute of Mental Health
Department of Psychiatry
University of British Columbia
430 - 5950 University Blvd
David Strangway Bldg, University of British Columbia
Vancouver, BC, V6T 1Z3
Canada
Phone: 1 604 362 5153
Fax: 1 604 827 3373
Email: feliciawetterlin@gmail.com

Abstract

Background: Due to the high prevalence of psychological disorders and the lack of access to care among Canadian youth, the development of accessible services is increasingly important. eMental Health is an expanding field that may help to meet this need through the provision of mental health care using technology.

Objective: The primary goals of the study are to explore youth experiences with traditional and online mental health resources, and to investigate youth expectations for mental health websites.

Methods: A Web-based survey containing quantitative and qualitative questions was delivered to youth aged 17-24 years. Participants were surveyed to evaluate their use of mental health resources as well as their preferences for various components of a potential mental health website.

Results: A total of 521 surveys were completed. Most participants (61.6%, 321/521) indicated that they had used the Internet to seek information or help for feelings they were experiencing. If they were going through a difficult time, 82.9% (432/521) of participants were either “somewhat likely” or “very likely” to use an information-based website and 76.8% (400/521) reported that they were either “somewhat unlikely” or “very unlikely” to visit social media websites for information or help-seeking purposes during this time. Most (87.7%, 458/521) participants rated their online privacy as very important. Descriptions of interventions and treatments was the most highly rated feature to have in a mental health-related website, with 91.9% (479/521) of participants regarding it as “important” or “very important”. When presented a select list of existing Canadian mental health-related websites, most participants had not accessed any of the sites. Of the few who had, the Canadian Mental Health Association website was the most accessed website (5.8%, 30/521). Other mental health-related websites were accessed by only 10.9% of the participants (57/521).

Conclusions: The findings suggest that despite interest in these tools, current eMental Health resources either do not meet the needs of or are not widely accessed by youth with mental health problems. In order to improve access to these resources for Canadian youth, Web-based platforms should provide information about mental health problems, support for these problems (peer and professional), and information about resources (self-help as well as ability to locate nearby resources), while protecting the privacy of the user. These findings will not only assist in the development of new mental health platforms but may also help improve existing ones.

XSL•FO
RenderX

(J Med Internet Res 2014;16(12):e293) doi:10.2196/jmir.3526

KEYWORDS

mental health services; online systems; adolescents; survey; Internet

Introduction

Canadians are heavy users of the Internet [1]. In 2011, 83% of Canada's population reported using online services [2], and in 2014, 99% of young Canadians reported having access to the Internet outside of school [3]. The number of individuals using the Internet to search for medical or health-related information increased from 57.9% in 2005 to 69.9% in 2009 [4]. With such a large part of Canada's population online, and with many young adults using the Internet for health-seeking purposes, it is critical to understand more about the health-seeking behaviors, experiences, and expectations of youth.

eMental Health is an expanding field and is defined by the Mental Health Commission of Canada as "the use of information and communication technologies to improve mental health" [5]. An increased use of the Internet in everyday life has contributed to a rise in self-help and Web-based therapies for various physical and psychological disorders [6,7].

Developing more accessible services has become increasingly important due to the high prevalence of mental disorders—particularly mood, substance use, and anxiety disorders—in Canada. According to the Canadian Mental Health Survey, 20% of Canadians will experience a mental illness during their lifetime, 11.3% will be diagnosed with major depressive disorder, and 8.7% with an anxiety disorder [8]. The National Comorbidity survey in the United States reports an even higher lifetime mental disorder prevalence of 44% [9].

Youth between the ages of 15 and 24 years have the highest prevalence of mental disorders, and yet almost one-third of Canadians seeking mental health care report that their needs are unmet or partially unmet [8]. In one study, only 42% of young adults aged 19-24 years with major depressive disorder had accessed any mental health services [10]. Research suggests that youth view current/traditional mental health services negatively, and that future efforts should center on the development of interconnected and empowering services as a means of increasing service use [11]. The Canadian Institutes of Health Research states that "advanced eHealth innovations feature integrated and sophisticated technology that enables interconnectivity at all levels of the health system" and that "eHealth empowers patients and encourages self-management" [12]. These facts suggest the high level of unmet needs can only be addressed by the addition of significant capacity. The need to develop Web-based mental health services is supported by the perceived acceptability of online therapy as a form of treatment [13], as well as its effectiveness in providing support for youth who have family members experiencing mental illness [14].

To guide the conceptualization and development of future effective, attractive, and youth-appropriate Web-based services, we studied current experiences and expectations of online health solutions among youth by exploring the following questions: (1) What sources of online information are young people most likely to use, and what do they find most accurate, when seeking help for mental health issues?, (2) What sources of online information do young people predict they would use if they had a mental health issue?, and (3) What experience do young people have in using mental health resources on the Internet?

Methods

Recruitment

The Bell Youth Impact Survey ran from September 2012 to March 2013 and garnered 521 complete responses. Recruitment of youth living in Canada between 17 and 24 years of age was accomplished through a variety of methods. Participants were recruited via targeted Facebook advertisements; Facebook and Twitter pages dedicated specifically to the survey; internal communiqués at the Centre for Student Involvement at the University of British Columbia, Student Communications Services, the Centre for Health Evaluation and Outcome Sciences, the Paid Participants Studies list; and through partner community organizations.

Survey Procedure

The survey was developed after a review of existing literature concerning youth self-reports of their mental health and resource use. Questions were created in consultation with a licensed psychiatrist. The survey was then administered to members of the research team as well as non-affiliated individuals for pre-testing before use. Approval for the survey was obtained from the Behavioral Research Ethics Board at the University of British Columbia. All participants, regardless of recruitment source, were directed to a website providing a brief description of the survey. From this website, youth interested in participating were prompted to email the researcher in order to receive a website link to the survey and a unique, single-use access code. Participants who followed the link were presented with a consent form and upon acceptance were redirected to the survey, which was hosted on FluidSurveys.

The first component of the survey consisted of the exclusion criteria. Any participant indicating that they were 16 years or younger, 25 years or older, or not currently living in Canada was redirected to the survey termination page and notified that they were not eligible to complete the survey.

The survey consisted of approximately 65 questions distributed over a maximum of 16 pages (depending on skip logic), and took between 15 and 20 minutes to complete. Participants were permitted to skip any question they were unwilling to answer and to discontinue the survey at any time. They were also able to review and change their answers during the course of the survey. Participants who completed the survey and provided an email address received a CAN $5 gift card to a retail store of their choice (Starbucks, Chapters, or Amazon).

XSL•FO
RenderX

Only completed questionnaires, in which all pages were visited (even if some were skipped), were included in the analysis. Participant identifiers were removed from the survey before data analysis.

Measures

The survey consisted of both quantitative and qualitative questions to assess: (1) demographics, (2) mental health literacy, (3) online and offline use of mental health resources, (4) online and offline opinions about current mental health resources, and (5) preferences regarding potential components of a mental health website. In this paper, we include items (3), (4), and (5).

Results

Participant Demographics

Out of the 521 participants, 76.6% were female (399/521) and the mean age was 20.68 years (SD 2.08). Most participants were East/Southeast Asian (44.0%, 229/521) or European/Caucasian (35.5%, 185/521). Most (61.6%, 321/521) participants indicated that they had used the Internet to seek information or help for the feelings they were experiencing. Participant location, ethnicity, and education status are presented in Table 1.

When asked what they were looking for on the Internet in the context of seeking mental health information (Table 2), 52.4% of participants (273/521) indicated that when using the Internet for mental health information-seeking purposes they were looking for information about symptoms and 47.4% (247/521) were looking for treatment options. A total of 2.5% participants (13/521) selected “other” as a response and provided their own answer, with common themes such as looking for personal testimonies from people with the same mental health problem, coping mechanisms for their family, counseling or crisis chats, self-help strategies, and ways in which to cope.

Table 1. Participant demographics (n=521).

Demographic		n (%)
Gender		
	Male	121 (23.2)
	Female	399 (76.6)
Current location		
	British Columbia	448 (86.0)
	Ontario	43 (8.3)
	Newfoundland	10 (1.9)
	Nova Scotia	6 (1.2)
	Alberta	3 (0.6)
	Manitoba	3 (0.6)
	Quebec	3 (0.6)
	Prince Edward Island	1 (0.2)
	Yukon	1 (0.2)
Student status		
	Currently in school (eg, high school or post-secondary education)	455 (87.3)
	Not currently in school	65 (12.5)
Ethnicity		
	East/Southeast Asian	229 (44.0)
	European/Caucasian	185 (35.5)
	South Asian	63 (12.2)
	West Asian	32 (6.1)
	Latin/Central/South American	15 (2.9)
	Aboriginal	8 (1.5)
	Caribbean	7 (1.3)
	Arab	5 (1.0)
	African	5 (1.0)
	Pacific Islander	3 (0.6)

XSL•FO
RenderX

Table 2. Web-based mental health information-seeking tendencies (n=521).

Features	n (%)
Information about symptoms	273 (52.4)
Information about treatment options	247 (47.4)
Information about prevalence rates	90 (17.3)
Web-based questionnaires or assessment tests	124 (23.8)
Peer support	82 (15.7)
A list of local resources	68 (13.1)
Other	13 (2.5)

Resource Use

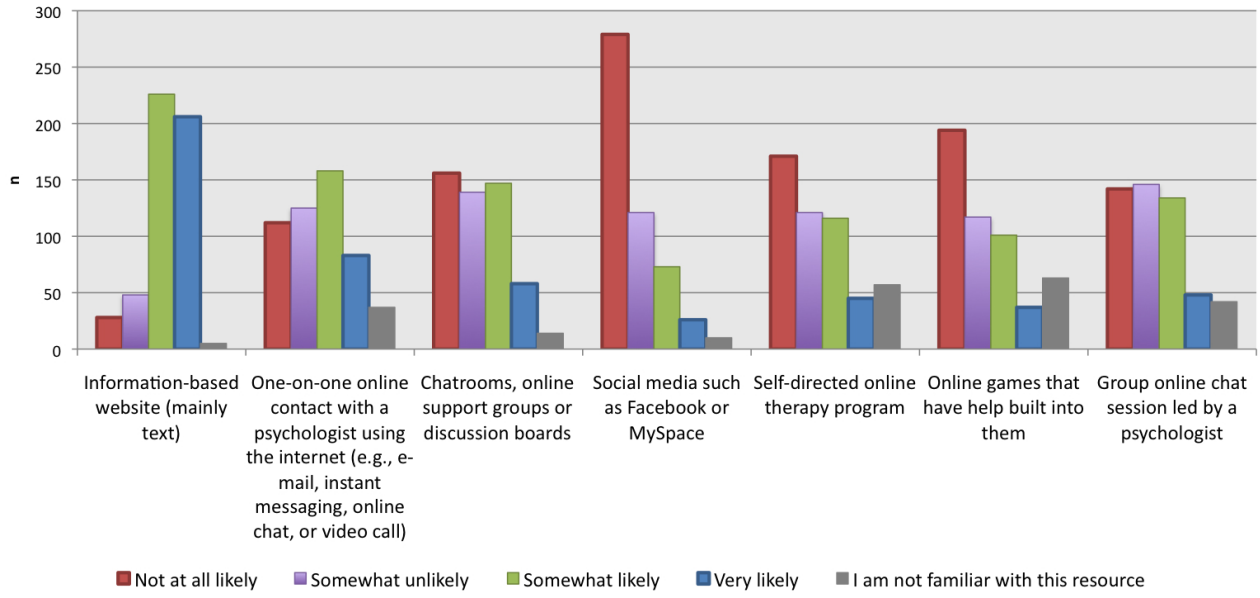
When asked about the likelihood of visiting certain types of Web-based resources during a difficult time in life (Figure 1), most (82.9%, 432/521) participants were either "somewhat likely" or "very likely" to use an information-based website with mainly text. Slightly more than half of the participants were "not at all likely" or "somewhat unlikely" to use online interactions such as a group online chat session led by a psychologist (55.3%, 288/521) and chat rooms/support groups/discussion boards (56.6%, 295/521). More than half (59.7%, 311/521) of participants stated that they were "not at all likely" or "somewhat unlikely" to use online games that have help built into them. The majority (76.8%, 400/521) reported that they were either "somewhat unlikely" or "very unlikely" to visit social media websites during this time.

When asked to rate the importance of human contact within a Web-based mental health resource (Figure 2), the majority of participants (83.9%, 437/521) listed contact with an online professional (eg, therapist or coach) as either somewhat important or very important. Peer support (81.8%, 426/521), family involvement (74.8%, 390/521), and friend involvement (77.1%, 402/521) were other popular responses listed as "somewhat important" or "very important" as contacts to have within a Web-based mental health resource.

Out of the 322 participants who indicated that they had used the Internet to look for information or help about what they were feeling, only 10.6% (34/321) said that they had used social media (eg, Facebook, MySpace) to obtain help with problems such as anxiety or depression.

Figure 3 shows the values and corresponding percentages of the responses given when participants were asked to rate the importance of certain features as components of a Web-based mental health resource. The most highly rated features (regarded as "important" or "very important") by participants were descriptions of interventions and treatments (91.9%, 479/521), evidence-based information and links to scientific or research articles (88.3%, 460/521), a resource list to access help "in your area" (86.7%, 452/521), self-guided Web-based interventions for anxiety and depression (84.1%, 438/521), self-help information and tools (82.2%, 428/521), quizzes and tools to help assess mood and behavior (80.0%, 417/521), pictures to help explain topics (75.8%, 395/521), and videos to help explain topics (72.3%, 377/521).

Figure 1. Likelihood of visiting Web-based resources during a difficult time in life.



XSL•FO
RenderX

Figure 2. Importance of contact with different supports in a Web-based mental health resource.

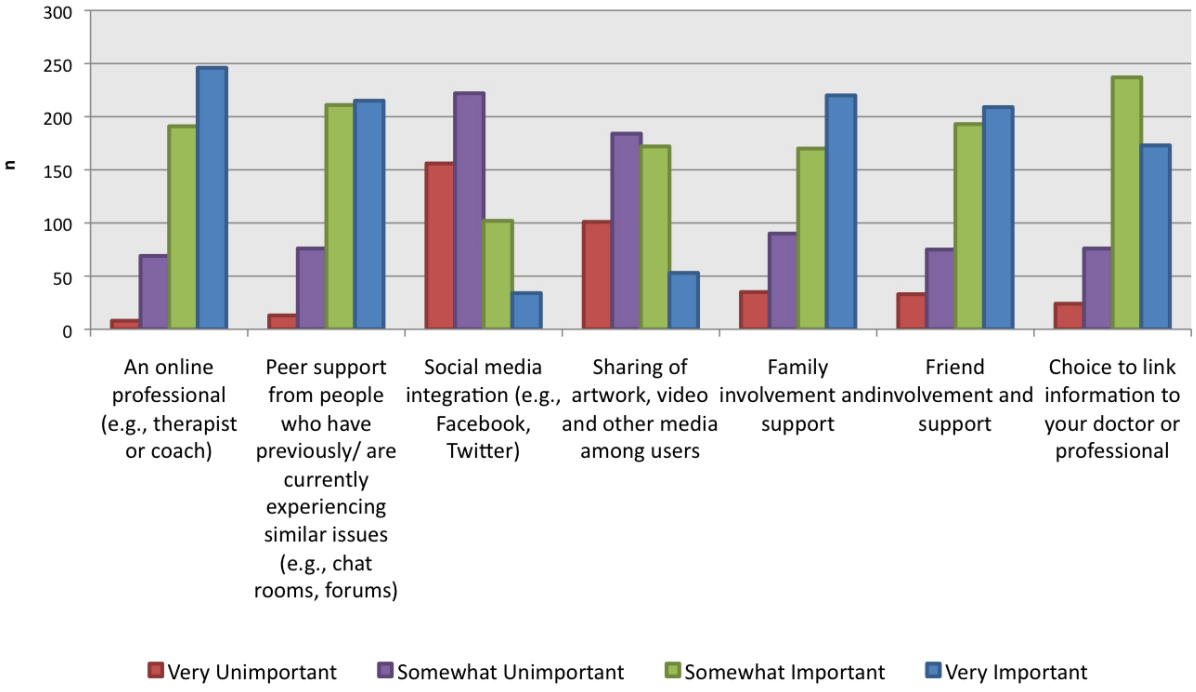
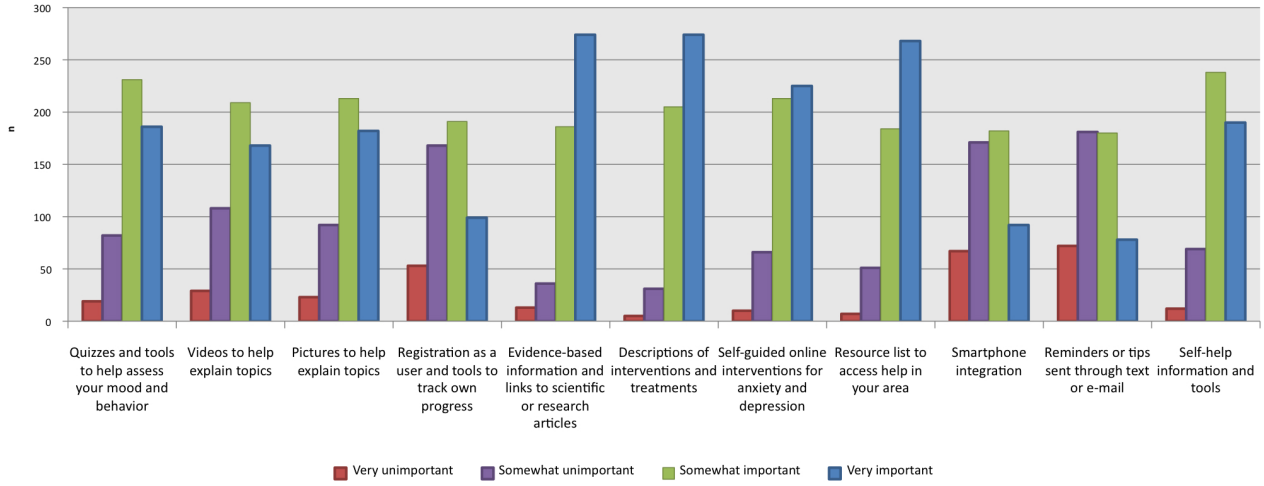


Figure 3. Importance of different features within a Web-based mental health resource.



Privacy

Most participants (87.7%, 458/521) rated their privacy as a user as “very important”.

Website Familiarity

Participant familiarity with current popular mental health-related websites was assessed by asking participants to indicate, from a list of Canadian websites related to mental health and well-being, which they had visited. Most participants had not visited any of the websites, but, of the minority who had, the Canadian Mental Health Association website was the most highly accessed Canadian mental health-related website (5.8%, 30/521). The remaining websites were only accessed by 10.9% of the participants (57/521).

Discussion

Principal Findings

With 93% of teens and young adults regularly online [15], Web-based mental health resources need to become a priority. Easy access to information and treatment options about mental health disorders can better facilitate early detection and treatment of psychological disorders in youth. Our findings show that the Internet is a key resource for youth between the ages of 17 and 24 years when searching for information on symptoms or when looking for help concerning their mental health. This focus on Web-based resources may be linked to the inadequacy of offline/traditional mental health services in providing information and services for all who seek it [16].

Participants rated the importance of online support in the form of online contact with a psychologist or mental health

XSL•FO
RenderX

professional, peer support, friend involvement, and family involvement as important features of a mental health-related website. This finding may suggest an underlying desire to seek human interaction as a part of an online community without interfering with daily life or compromising privacy.

However, when asked to rate their personal likelihood of using services designed to provide support (such as group online chats with a psychologist, chat rooms, and support groups), the majority of participants stated that they were unlikely to use these services. This discrepancy in results between Figures 1 and 2 may be influenced by our sample, which included both those with mental health issues and those without. For example, participants without mental health problems may have underestimated what features they would use if they were going through a difficult time in their life. Alternatively, this difference in responses may be explained by the psychological phenomenon—pluralistic ignorance—in which people tend to collectively assume that their opinion of a norm or belief is the minority and that the majority thinks differently than they do [17]. In the words of Krech and Crutchfield, "no one believes, but everyone thinks that everyone believes" [18]. In this instance, the average person did not personally need online support, but believes that others do. If this is the case, the response rate of Figure 2 may be an overestimate of the importance of contact with different supports in Web-based mental health resources for youth. Other research has found similar trends in hypothetical information-seeking and actual usage of Web-based sources, with a higher percentage of participants indicating that they were likely to use a certain feature than the percentage of participants who had actually used the feature [19].

Youth do not appear to use social networking sites when looking for help or support for problems such as anxiety or depression. Most participants rated their privacy on the Internet as very important, which may explain the lack of social networking site use for mental health help or support. This avoidance of social networking sites may also be due to the stigma associated with mental illness [20] and the public nature of such sites. Because of the reported tendency for youth to not use social networks for mental health concerns, it may be important for Web-based platforms to discuss mental health problems in a way that fosters a sense of community while ensuring anonymity for the user. This could help youth with mental health problems connect with similar peers in a safe environment.

Most features (Figure 3) in the survey were rated as being important components of a mental health-related website, suggesting that a variety of features are likely to meet the needs of the youth population. A mental health website would need to have information from credible and trustworthy sources, including descriptions of interventions and treatments. This finding is consistent with previous research [21], which suggests that the current adolescent generation faces the challenge of finding reliable information online.

Of the currently available mental health-related websites we asked about, very few of the participants had accessed any at all. Reasons for this lack of access are unclear. It may be due to lack of awareness of existing mental health-related websites as a result of advertising (ie, perhaps these websites are helpful, but are not advertised and therefore not recognized or accessed by participants), or perhaps they are inadequate and there are no sites that meet users' needs. Alternatively, in our study, we may have inquired about the "wrong" websites (ie, websites that do not provide the needed services and are therefore not used by young people searching for mental health resources) but other sites are used and seen to be useful. This finding is supported by the systematic review by Krauer et al [22], which analyzed help-seeking behavior among young people and found that current Web-based mental health resources do not contribute significantly to help-seeking behaviors. The review, however, does state that young people are inclined to use Web-based resources. Rickwood et al have further explored the subject in a review that underlined the need for mental-health help to be easily accessible and youth-friendly [23]. Help-seeking behavior in youth might therefore be enhanced by aligning Web-based mental health resources with the preferences of youth as identified in this paper.

Limitations

The findings of this study are limited to youth between the ages of 17-24 years living in Canada, the majority of whom are pursuing a post-secondary education. Females are more likely to seek out and engage with services, so our results may not generalize to the Canadian youth population as a whole [23]. Most participants were living in British Columbia and reflected the general make-up of the province [24], and it is reasonable to expect similar preferences from youth across different provinces. Although the sample largely consisted of European/Caucasian and East/South East Asian youth, it should be noted that the survey was designed in such a way that participants could identify as more than one ethnicity. Therefore, any generalizations between ethnicity and Web-based mental health service use and expectations should be made with caution. Another limitation is that the study relied on self-reported data to determine eMental Health resource use and the expectations of future Web-based resources. Because some of the questions allowed the participants to select more than one response, it was not possible to determine rank of choices and therefore popularity of the features relative to one another. Last, this survey also asked participants about hypothetical situations, emotions, and Web-based mental health resources, which may not be reflective of actual behavior.

Conclusions

These findings suggest that the next step in Web-based mental health care is to ensure that online platforms (not social media-based) provide information about mental health problems, support for these problems (peer and professional), and resources (self-help as well as information on nearby resources), while protecting the privacy of the user. These findings will not only assist in the development of new mental health platforms, they may also help improve existing ones.

XSL•FO
RenderX

Acknowledgments

This study was funded by a grant from Bell's "Let's Talk" Initiative.

Conflicts of Interest

The authors disclose affiliation with the freely available youth eMental health service "WalkAlong".

References

1. Canadian Internet Registration Authority. 2013. CIRA Factbook URL: http://www.cira.ca/factbook/2013/index.html [accessed 2014-03-03] [WebCite Cache ID 6No1w0ms8]
2. Statistics Canada. 2012. Individual Internet use and e-commerce URL: http://www.statcan.gc.ca/daily-quotidien/131028/dq131028a-eng.pdf [accessed 2013-11-21] [WebCite Cache ID 6LIuagwfN]
3. Steeves, Valerie. MediaSmarts. 2014. Young Canadians in a wired world, phase III: life online URL: http://mediasmarts.ca/sites/mediasmarts/files/pdfs/publication-report/full/YCWWIII_Life_Online_FullReport_2.pdf [accessed 2014-10-10] [WebCite Cache ID 6TDzjJ6uL]
4. Statistics Canada. 2010. Individual and household Internet use URL: http://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/comm29a-eng.htm [accessed 2013-11-20] [WebCite Cache ID 6LIuiTHO3]
5. Mental Health Commission of Canada. 2012. Issue: e-Mental health URL: http://www.mentalhealthcommission.ca/English/issues/e-mental-health?routetoken=a811a0064f3e820a2b5a21aadec14b4e&terminitial=131 [accessed 2013-11-21] [WebCite Cache ID 6LIuslbyl]
6. Barak A, Hen L, Boniel-Nissim M, Shapira N. A comprehensive review and a meta-analysis of the effectiveness of Internet-based psychotherapeutic interventions. Journal of Technology in Human Services 2008 Jul 03;26(2-4):109-160. [doi: 10.1080/15228830802094429]
7. Stinson J, Wilson R, Gill N, Yamada J, Holt J. A systematic review of internet-based self-management interventions for youth with health conditions. J Pediatr Psychol 2009 Jun;34(5):495-510 [FREE Full text] [doi: 10.1093/jpepsy/jsn115] [Medline: 19029142]
8. Statistics Canada. 2013. Mental and substance use disorders in Canada URL: http://www.statcan.gc.ca/pub/82-624-x/2013001/article/11855-eng.pdf [accessed 2013-11-21] [WebCite Cache ID 6LIuydERb]
9. 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]
10. Cheung AH, Dewa CS. Mental health service use among adolescents and young adults with major depressive disorder and suicidality. Can J Psychiatry 2007 Apr;52(4):228-232. [Medline: 17500303]
11. Marcus MA, Westra HA, Eastwood JD, Barnes KL, Mobilizing Minds Research Group. What are young adults saying about mental health? An analysis of Internet blogs. J Med Internet Res 2012 Jan;14(1):e17 [FREE Full text] [doi: 10.2196/jmir.1868] [Medline: 22569642]
12. Canadian Institutes of Health Research. 2013. About e-Health innovations URL: http://www.cihr-irsc.gc.ca/e/47345.html [accessed 2014-03-05] [WebCite Cache ID 6NqxUicR3]
13. Schneider J, Sarrami Foroushani P, Grime P, Thornicroft G. Acceptability of online self-help to people with depression: users' views of MoodGYM versus informational websites. J Med Internet Res 2014 Mar;16(3):e90 [FREE Full text] [doi: 10.2196/jmir.2871] [Medline: 24681717]
14. Drost LM, Cuijpers P, Schippers GM. Developing an interactive website for adolescents with a mentally ill family member. Clin Child Psychol Psychiatry 2011 Jul;16(3):351-364. [doi: 10.1177/1359104510366281] [Medline: 20980365]
15. Lenhart A, Purcell K, Smith A, Zickuhr K. Pew Internet & American Life Project. 2010 Feb 03. Social media & mobile internet use among teens and young adults URL: http://www.pewinternet.org/files/old-media//Files/Reports/2010/PIP_Social_Media_and_Young_Adults_Report_Final_with_toplines.pdf [accessed 2014-12-01] [WebCite Cache ID 6UVNyvNGv]
16. Harris Interactive. Rochester, New York; 2008. Mental health treatment: it's commonly accepted yet not so easy to obtain or understand URL: http://www.harrisinteractive.com/NewsRoom/PressReleases/tabid/446/ctl/ReadCustom%20Default/mid/1506/ArticleId/311/Default.aspx [accessed 2013-11-21] [WebCite Cache ID 6LIv4Ekw4]
17. Katz D, Allport FH, Jeness MB. Students' Attitudes: A report of the Syracuse University reaction study. Syracuse, New York: The Craftsman Press; 1931.
18. Krech D, Crutchfield RS, Ballachey EL. Theory and problems of social psychology. New York: McGraw-Hill; 1948.
19. Neal DM, Campbell AJ, Williams LY, Liu Y, Nussbaumer D. "I did not realize so many options are available": Cognitive authority, emerging adults, and e-mental health. Library & Information Science Research 2011 Jan;33(1):25-33. [doi: 10.1016/j.lisr.2010.07.015]
20. Eisenberg D, Downs MF, Golberstein E, Zivin K. Stigma and help seeking for mental health among college students. Med Care Res Rev 2009 Oct;66(5):522-541. [doi: 10.1177/1077558709335173] [Medline: 19454625]

XSL•FO
RenderX

21. Rieh SY, Hilligoss B. College students' credibility: judgments in the information-seeking process. In: Digital media, youth, and credibility. Cambridge, Mass: MIT Press; 2008:49-71.
22. Kauer SD, Mangan C, Sanci L. Do online mental health services improve help-seeking for young people? A systematic review. J Med Internet Res 2014 Mar;16(3):e66 [FREE Full text] [doi: 10.2196/jmir.3103] [Medline: 24594922]
23. Rickwood D, Deane FP, Wilson CJ, Ciarrochi J. Young people's help-seeking for mental health problems. Australian E-Journal for the Advancement of Mental Health 2005;4(3):1-34 [FREE Full text]
24. Statistics Canada. 2011 National Household Survey: Data tables URL: http://www12.statcan.gc.ca/nhs-enm/2011/dp-pd/dt-td/Index-eng.cfm [accessed 2014-10-10] [WebCite Cache ID 6TE03Peks]

Edited by G Eysenbach; submitted 09.05.14; peer-reviewed by S Kauer, C Connor, G Schippers; comments to author 17.08.14; revised version received 10.10.14; accepted 06.11.14; published 17.12.14.

Please cite as:
Wetterlin FM, Mar MY, Neilson EK, Werker GR, Krausz M
eMental Health Experiences and Expectations: A Survey of Youths' Web-Based Resource Preferences in Canada
J Med Internet Res 2014;16(12):e293
URL: http://www.jmir.org/2014/12/e293/
doi:10.2196/jmir.3526
PMID:25519847

©Felicia M Wetterlin, Marissa Y Mar, Erika K Neilson, Gregory R Werker, Michael Krausz. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 17.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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

Cumulative Query Method for Influenza Surveillance Using Search Engine Data

Dong-Woo Seo[1], MD, PhD; Min-Woo Jo[2], MD, PhD; Chang Hwan Sohn[1], MD; Soo-Yong Shin[3], PhD; JaeHo Lee[1,3,4,5], MD, PhD; Maengsoo Yu[6]; Won Young Kim[1], MD, PhD; Kyoung Soo Lim[1], MD, PhD; Sang-Il Lee[2], MD, PhD

[1]Asan Medical Center, Department of Emergency Medicine, University of Ulsan, College of Medicine, Seoul, Republic Of Korea

[2]Asan Medical Center, Department of Preventive Medicine, University of Ulsan, College of Medicine, Seoul, Republic Of Korea

[3]Asan Medical Center, Department of Biomedical Informatics, Seoul, Republic Of Korea

[4]Brigham and Women's Hospital, Division of General Medicine and Primary Care, Boston, MA, United States

[5]Harvard Medical School, Boston, MA, United States

[6]Daum Communications, Search Development Unit, Seoul, Republic Of Korea

Corresponding Author:
Min-Woo Jo, MD, PhD
Asan Medical Center
Department of Preventive Medicine
University of Ulsan, College of Medicine
88, Olympic-Ro 43-Gil, Songpa-gu
Seoul, 138-769
Republic Of Korea
Phone: 82 2 3010 3350
Fax: 82 2 3010 3356
Email: leiseo@hanmail.net

Abstract

Background: Internet search queries have become an important data source in syndromic surveillance system. However, there is currently no syndromic surveillance system using Internet search query data in South Korea.

Objectives: The objective of this study was to examine correlations between our cumulative query method and national influenza surveillance data.

Methods: Our study was based on the local search engine, Daum (approximately 25% market share), and influenza-like illness (ILI) data from the Korea Centers for Disease Control and Prevention. A quota sampling survey was conducted with 200 participants to obtain popular queries. We divided the study period into two sets: Set 1 (the 2009/10 epidemiological year for development set 1 and 2010/11 for validation set 1) and Set 2 (2010/11 for development Set 2 and 2011/12 for validation Set 2). Pearson's correlation coefficients were calculated between the Daum data and the ILI data for the development set. We selected the combined queries for which the correlation coefficients were .7 or higher and listed them in descending order. Then, we created a cumulative query method *n* representing the number of cumulative combined queries in descending order of the correlation coefficient.

Results: In validation set 1, 13 cumulative query methods were applied, and 8 had higher correlation coefficients (min=.916, max=.943) than that of the highest single combined query. Further, 11 of 13 cumulative query methods had an *r* value of ≥.7, but 4 of 13 combined queries had an *r* value of ≥.7. In validation set 2, 8 of 15 cumulative query methods showed higher correlation coefficients (min=.975, max=.987) than that of the highest single combined query. All 15 cumulative query methods had an *r* value of ≥.7, but 6 of 15 combined queries had an *r* value of ≥.7.

Conclusions: Cumulative query method showed relatively higher correlation with national influenza surveillance data than combined queries in the development and validation set.

(J Med Internet Res 2014;16(12):e289) doi:10.2196/jmir.3680

KEYWORDS

syndromic surveillance system; influenza; influenza-like illness; Google Flu Trends; Internet search; query

XSL•FO
RenderX

Introduction

Syndromic surveillance may alert public health care providers in the early phases of an outbreak, allowing them to decrease morbidity and mortality resulting from the outbreak [1-5]. Syndromic surveillance is defined as the real-time or near real-time collection, analysis, interpretation, and dissemination of health-related data to enable the early identification of the impact of potential human or veterinary public health threats that require effective public health action [1,3]. The 2009 H1N1 influenza pandemic highlighted the need for syndromic surveillance to inform policy and plan for effective responses.

Because conventional syndromic surveillance of indicators such as influenza-like illness (ILI) depends on case reporting to report disease activity, time delays in reporting and case confirmation can interfere with the early detection of outbreaks or increases in influenza cases in the community. Thus, researchers have been investigating alternative data sources for the detection of outbreaks. For example, over-the-counter sales of medications and school absenteeism data have been used for earlier detection of outbreaks [6-12].

Internet search queries have become an important data source in recent years [13-22]. Internet search engines allow billions of people to have instant access to a vast amount of information online. New syndromic surveillance sources, such as Google Flu Trends (GFT), provide the potential to identify influenza outbreaks in real time [23]. Several studies have reported that GFT is highly correlated with conventional ILI surveillance data [23-28]. GFT has now been applied in many countries, but neither GFT nor other search query-based tools for disease surveillance are available in South Korea [24,27-30]. Generally, Google's market share is dominant in the countries where GFT is available [24-29,31], but not in South Korea [32]. Studies using Google Trends for influenza surveillance show that it can be used as a complementary source of data but that its performance is insufficient for use as a model for prediction [33,34]. It is difficult to find queries that show high correlations for consecutive years because Internet searching behavior may change over time [33,34]. To reduce the effects from changes in search queries, we used a combination of queries and cumulation of combined queries from the search engine Daum. Daum is the second largest Web portal service provider in daily visits in South Korea (approximately 25% of the market share) [32,35]. Daum offers many Internet services to Web users, including email, messaging service, forums, shopping, and news. The main language is Korean.

In South Korea, influenza is generally seasonal, with most activity occurring during winter. The 2009/10 epidemiological year, called the Influenza A (H1N1) pandemic period, was an exceptional situation (see Multimedia Appendix 1). The primary objective of this study was to examine correlations between our cumulative query method and national influenza surveillance data.

Methods

Source of Data

Study Period

The study period was September 6, 2009 (week 36), through September 1, 2012 (week 34)—156 weeks of data for 3 consecutive epidemiological years. We divided the study period into two sets: Set 1 (the 2009/10 epidemiological year for development set 1 and 2010/11 for validation set 1) and Set 2 (2010/11 for development set 2 and 2011/12 for validation set 2).

Collection of Influenza-Like Illness Data

We collected the ILI data from the Korea Centers for Disease Control and Prevention (KCDC) as a gold standard. KCDC ILI data were available from the KCDC website; we downloaded the ILI data for the study period from this site [36]. A KCDC case of ILI was defined as a person with a fever of 38°C with a cough and/or a sore throat [36]. ILI surveillance consisted of 850 sentinel clinics in South Korea, and the clinics reported weekly percentages of outpatients who met the case definition of ILI [36].

Survey for Obtaining Queries

To obtain population search queries related to influenza, we conducted a survey from quota sampling based on sex and age in September 2012. The quotas were based on address of resident registry, age, and sex. There were five quota groups by age: 20-29 years, 30-39 years, 40-49 years, 50-59 years, 60 years or older. Half of each quota group were female. We randomly selected the addresses from the residence registry in Seoul, and then if interviewees living at the address of residence registry met the criteria, we included the oldest interviewee. We then conducted face-to-face interviews. The survey included searching history for influenza and typed queries. The survey was performed anonymously. A KCDC definition of ILI was a person with a fever (발열 in Korean) of 38°C with a cough (기침) and/or a sore throat (인후통). These three queries from the definitions of ILI were included in the queries for the following operations, regardless of the survey result. In the case of queries originally submitted in English only, we translated them to Korean and added them as new queries.

Combination of Queries

We believe that people typically search for things of interest on the Internet using one or more queries at a time. To reflect people's searching behavior and include as many queries as possible, we used a combination of queries. Queries from the survey results and the definition of KCDC ILI were divided into groups as follows: query group 1 consisted of queries specific to influenza (eg, "H1N1", "Influenza"), and query group 2 contained queries not specific to influenza (eg, "Treatment", "Symptom"). Then, we combined query groups 1 and 2. Combined queries consisted of query group 1 alone and a combination of query groups 1 and 2 (eg, "H1N1", "H1N1 Treatment", "H1N1 Symptom", "Influenza", "Influenza Treatment", "Influenza Symptom").

XSL•FO
RenderX

Collection of Data from Search Engine

We sent the combined queries and the queries that belonged to query group 1 (because these queries were searchable by themselves) to Daum and received proportional data in weekly form. Proportional data for these combined queries were extracted from the Daum search engine during development sets 1 and 2. Proportional data from the Daum search engine were calculated by dividing the number of each combined query by the total number of search queries for 1 week.

Data Analysis

Creating Cumulative Query Methods and Data Analysis

Pearson's correlation coefficients were calculated between the Daum data for the combined queries and the KCDC ILI data in development sets 1 and 2. We selected the combined queries for which the correlation coefficients were .7 or higher and listed them in descending order. To see the change of correlation coefficients over time, we also calculated correlation coefficients of the combined queries in subsequent epidemiological years. We then created a cumulative query method *n* representing the number of cumulative combined queries in descending order of the correlation coefficient. For example, cumulative query method 4 consisted of a summation of the proportional data from the 1st, 2nd, 3rd, and 4th highest combined queries on the correlation coefficient list. In validation sets 1 and 2, Pearson's correlation coefficients were calculated between the cumulative query method *n* and the KCDC ILI data. Specifically in validation set 2, we analyzed the cumulative query methods from development set 2 as well as development set 1. Useful cumulative query methods in the validation sets were defined as having higher correlation coefficienst than the highest correlation coefficient of a single combined query in the same development set. Analysis was performed using IBM SPSS Statistics software, version 20. Significance was set at $P<.05$.

Institutional Review Board

This study was approved by the Institutional Review Board of Asan Medical Center (Seoul, Korea).

Results

Survey for Obtaining Queries

We contacted 322 people and included 200 participants older than 20 years who lived in Seoul, Korea. Over a quarter (56/200, 28%) answered "Yes" to the question of searching history for influenza and provided search queries (Table 1).

XSL•FO
RenderX

Table 1. Results of the survey.

Raw data	English translation	Frequency (%)
신종	New	1 (1.8)
신종플루	New flu[a]	23 (41.1)
신종플루 증상	New flu symptom	1 (1.8)
신종플루 증세	New flu sign	1 (1.8)
신종플루, 독감	New flu, bad cold	2 (3.6)
신종플루, 목아픔	New flu, neck pain	1 (1.8)
신종플루, 백신, Tamiflu	New flu, vaccine, Tamiflu (English)[b]	1 (1.8)
신종플루, 신플 증상	New flu, new flu (abbr.)[c] symptom	1 (1.8)
신종플루, 인플루엔자, H1N1, PCR	New flu, influenza, H1N1 (English)[b], PCR[d] (English)[b]	1 (1.8)
신종플루, 조류독감	New flu, bird flu	1 (1.8)
신종플루의 치료, 합병증	New flu, treatment, complication	1 (1.8)
신종플루증상	New flu symptom	1 (1.8)
신종플루증세, 예방, 마스크	New flu sign, prevention, mask	1 (1.8)
신플증상	New flu (abbr.)[c] symptom	1 (1.8)
열, 기침	Fever, cough	1 (1.8)
유행성독감, influenza	Epidemic bad cold, influenza (English)[b]	1 (1.8)
인플루엔자	Influenza	7 (12.5)
인플루엔자, 신종독감, 신종 플루	Influenza, new bad cold, new flu	1 (1.8)
인플루엔자, 조류독감	Influenza, bird flu	1 (1.8)
인플루엔자, 조류독감, 돼지독감, 신종플루	Influenza, bird flu, swine flu, new flu	1 (1.8)
조류독감	Bird flu	5 (8.9)
조류독감, 사망	Bird flu, decease	1 (1.8)
증상, 목통증	Symptom, throat pain	1 (1.8)
Total		200 (100.0)

[a]Since the Influenza A (H1N1) pandemic period, media began to use “New flu (신종플루)” to distinguish the H1N1 influenza and previous influenzas in Korea. In 2010, KCDC announced that the official term was “Influenza (인플루엔자)”. But “New flu (신종플루)” and “Bad cold (독감)” are still more popular terms than “Flu (플루)” or “Influenza (인플루엔자)” in Korea. “Bad cold (독감)” in Korean has two meanings: one is influenza and the other, a severe common cold.

[b]The query was originally submitted in English.

[c]Abbreviation: “New flu (abbr.) (신플)” is the abbreviation of “New flu (신종플루)” in Korean.

[d]PCR: polymerase chain reaction.

Combination of Queries From the Survey

Query group 1 contained 14 queries that were specific to influenza, and query group 2 had 14 queries that were not specific to influenza (Table 2). A total of 210 combined queries were submitted to Daum. Full data of combined queries are presented in Multimedia Appendix 2.

XSL•FO
RenderX

Table 2. Query groups 1 and 2 from the survey results and the KCDC definition of ILI[a].

Query group 1	Query group 2
Flu	Vaccine
New flu	Prevention
New flu (abbr.)[b]	Mask
Influenza	Symptom
Influenza (English)[c]	Sign
New influenza	Cough
Bad cold[d]	Fever
New bad cold	Neck pain
Epidemic bad cold	Sore throat
H1N1 (English)[c]	Throat pain
Bird flu	PCR (English)[c,e]
Swine flu	Treatment
Tamiflu	Complication
Tamiflu (English)[c]	Decease

[a]Query group 1 consisted of queries specific to or related to influenza. Query group 2 contained queries not specific to influenza.

[b]Abbreviation.

[c]The query was originally submitted in English.

[d]"Bad cold (독감)" in Korean has two meanings: one is influenza and the other, a severe common cold. "Flu" in query group 1 is "플루" which is the English pronunciation written in Korean. In Korea, "Bad cold (독감)" is a more popular term than "Flu (플루)" or "Influenza (인플루엔자)".

[e]PCR: polymerase chain reaction.

Collection of Data From Search Engine

Correlation analysis was performed between the Daum data for combined queries and the KCDC ILI data in development sets 1 and 2 (Table 3). In development set 1, "New flu (abbr.)" had the highest correlation coefficient (r=.894, P<.001), and 13 combined queries had correlation coefficient r values of ≥.7. Among these 13 combined queries, the number of the combined queries that had correlation coefficient r values of ≥.7 was reduced to 4 in validation set 1 and to 2 in validation set 2. In development set 2, "Bad cold + Symptom" had the highest correlation coefficient (r=.969, P<.001), and a total of 15 combined queries had an r value of ≥.7. Among these 15 combined queries, the number of the combined queries that had correlation coefficient r values of ≥.7 was reduced to 6 in validation set 2. Only "Tamiflu" and "New flu + Symptom" showed correlation coefficients r values of ≥.7 for 3 consecutive years (Figure 1). The change of correlation coefficients for all combined queries over time are presented in Multimedia Appendix 2.

XSL•FO
RenderX

Table 3. Correlation analysis between the Daum data for combined queries and the KCDC ILI data in development sets 1 and 2.

Order	Combined query	Correlation coefficient			Combined query	Correlation coefficient	
		Development set 1 (2009/10)	Validation set 1 (2010/11)	Validation set 2 (2011/12)		Development set 2 (2010/11)	Validation set 2 (2011/12)
1	New flu (abbr.)[a]	.894[b]	.622[b]	[c]	Bad cold + Symptom	.969[b]	.981[b]
2	Flu + Vaccine	.871[b]	-.062[d]	-.157[e]	New flu + Treatment	.951[b]	.616[b]
3	New flu + Cough	.849[b]	.930[b]	.291[b]	New flu + Cough	.930[b]	.291[b]
4	New flu + Fever	.814[b]	.591[b]	.460[b]	New flu + Sign	.919[b]	.684[b]
5	Tamiflu + Vaccine	.805[b]	-.062[c]	[c]	Tamiflu	.904[b]	.981[b]
6	Tamiflu + Symptom	.800[b]	[c]	[c]	New influenza + Symptom	.896[b]	.650[b]
7	Flu + Symptom	.799[b]	.815[b]	.416[b]	Bad cold + Treatment	.887[b]	.814[b]
8	H1N1 + Symptom	.791[b]	[c]	[c]	Swine flu + Symptom	.877[b]	.005[e]
9	New flu + Sore throat	.738[b]	.504[b]	[c]	New flu + Symptom	.836[b]	.936[b]
10	New flu (abbr.)[a] + Vaccine	.713[b]	[c]	[c]	Flu + Symptom	.815[b]	.416[b]
11	New flu + Symptom	.709[b]	.836[b]	.936[b]	Influenza + Symptom	.813[b]	.782[b]
12	Tamiflu	.703[b]	.904[b]	.981[b]	Influenza (English)[g]	.762[b]	.751[b]
13	Tamiflu (English)[g]	.700b[b,h]	.523[b]	.286[b]	New influenza	.748[b]	.503[b]
14					Bird flu + Symptom	.747[b]	.005[f]
15					Bird flu	.709[b,h]	.136[i]

[a]abbr.: abbreviation

[b]$P<.05$.

[c]Correlation cannot be computed because it has a constant value in that period (see Multimedia Appendix 2).

[d]$P=.66$.

[e]$P=.27$.

[f]$P=.98$.

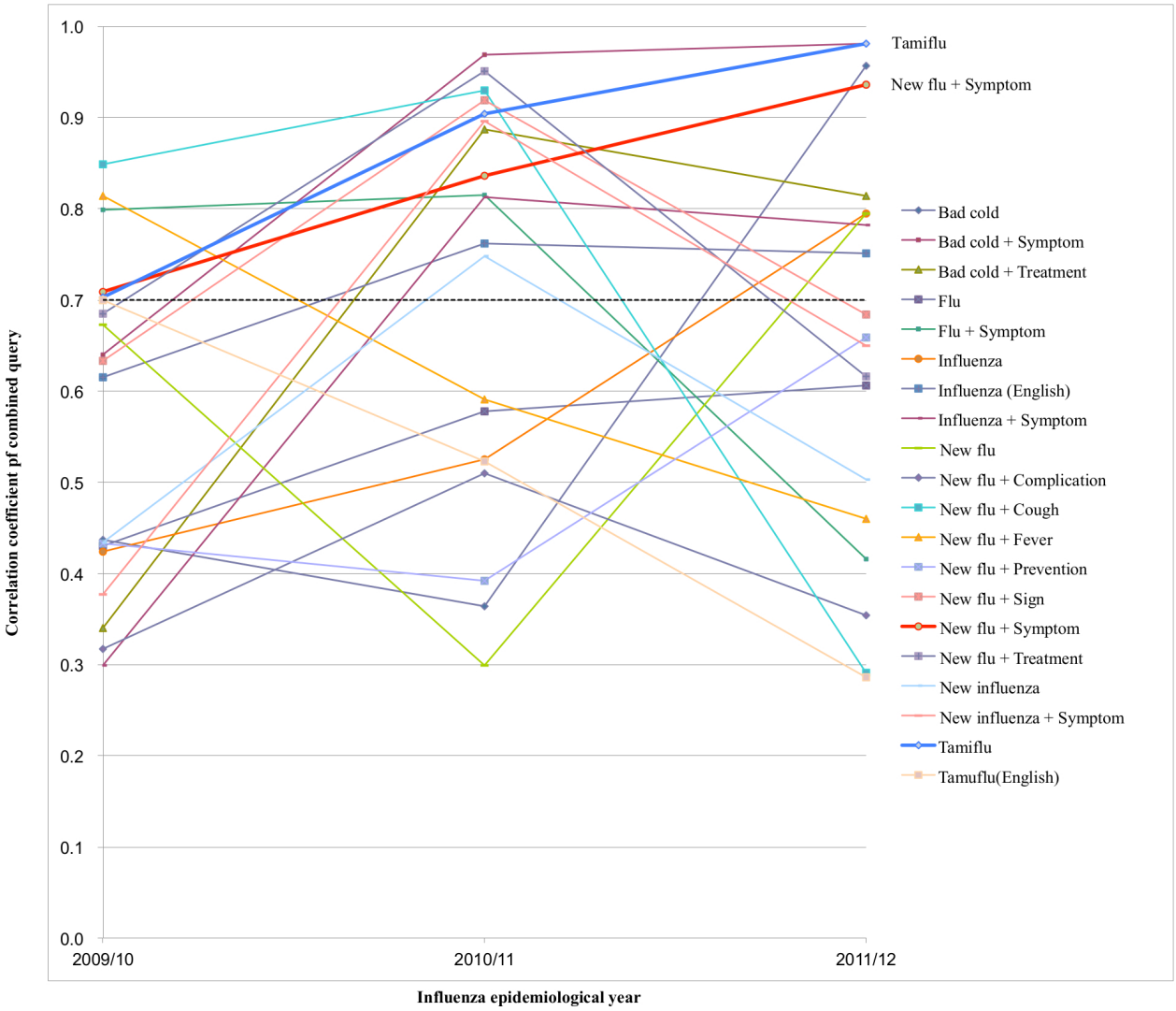
[g]The query was originally submitted in English.

[h]We selected the combined queries for which the correlation coefficients were ≥.7 and listed them in descending order.

[i]$P=.34$.

XSL•FO
RenderX

Figure 1. Plot of combined queries that consecutively show correlation coefficient (P<.05) (only "Tamiflu" and "New flu + Symptom" showed r values greater than .7 for 3 consecutive years).



Creating Cumulative Query Methods

A total of 13 cumulative query methods were created in development set 1 (see Table 4). In validation set 1, cumulative query methods 7, 8, 9, and 10 showed the highest correlation coefficients (r=.943, P<.001; see Multimedia Appendix 3). Eight of the 13 cumulative query methods were useful, which was defined as having higher correlation coefficients than the highest correlation coefficient of a single combined query in development set 1 (min=.916, max=.943). But only three of the cumulative query methods from development set 1 were useful in validation set 2 (min=.935, max=.953). The correlation did not increase by adding queries in cumulative query method 5, 6, 8, 9, 10, and 13 in validation set 1. In validation set 2, cumulative query method 5 from development set 2 had the highest correlation coefficient (r=.987, P<.001; see Figure 2 and Multimedia Appendix 3). Eight of the 15 cumulative query methods from development set 2 were useful (min=.975, max=.987). The correlation did not increase by adding queries in cumulative query method 3, 4, 7, 8, 10, 12, and 14 in validation set 2. Scatter plots between the KCDC ILI and other useful cumulative query methods are presented in Multimedia Appendix 4. Cumulative query methods for influenza virologic data are presented in Multimedia Appendix 5.

In each development set, cumulative query methods had a higher correlation coefficient than combined queries (see Tables 5 and 6). After 1 year, 11 of 13 cumulative query methods had an r value of ≥.7, but 4 of 13 combined queries had an r value of ≥.7 in validation set 1 (see Table 5 and Figure 3). All 15 cumulative query methods had an r value of ≥.7, but 6 of 15 combined queries had an r value of ≥.7 in validation set 2 (see Table 6 and Figure 4).

XSL•FO
RenderX

Table 4. Correlation coefficients of cumulative query method *n* in each validation set[a].

Cumulative query method	Correlation coefficient in validation set 1	Correlation coefficient in validation set 2 from development set 1	Correlation coefficient in validation set 2 from development set 2
1	.622[b]	[c]	.981[b,d]
2	.183[e]	-.157[f]	.975[b,d]
3	.916[b,d]	.092[g]	.975[b,d]
4	.933[b,d]	.467[b]	.975[b,d]
5	.933[b,d]	.467[b]	.987[b,d]
6	.933[b,d]	.467[b]	.986[b,d]
7	.943[b,d]	.486[b]	.986[b,d]
8	.943[b,d]	.486[b]	.986[b,d]
9	.943[b,d]	.486[b]	.968[b]
10	.943[b,d]	.486[b]	.968[b]
11	.838[b]	.935[b,d]	.965[b]
12	.841[b]	.953[b,d]	.965[b]
13	.841[b]	.953[b,d]	.964[b]
14	Not applicable	Not applicable	.964[b]
15	Not applicable	Not applicable	.780[b]

[a]We selected the combined queries for which the correlation coefficients were ≥.7 and listed them in descending order. We then created a cumulative query method *n* representing the number of cumulative combined queries in descending order of the correlation coefficients.

[b]$P<.05$.

[c]Correlation of cumulative query method 1 in validation set 2 from development set 1 cannot be computed because it has a constant value in that period (see Multimedia Appendix 2).

[d]Useful cumulative query method in the validation set was defined as having higher correlation coefficient than the highest correlation coefficient of a single combined query in the same development set.

[e]$P=.20$.

[f]$P=.27$.

[g]$P=.52$.

XSL•FO
RenderX

Figure 2. Scatter plot between the KCDC ILI and cumulative query model 5 in validation set 2.

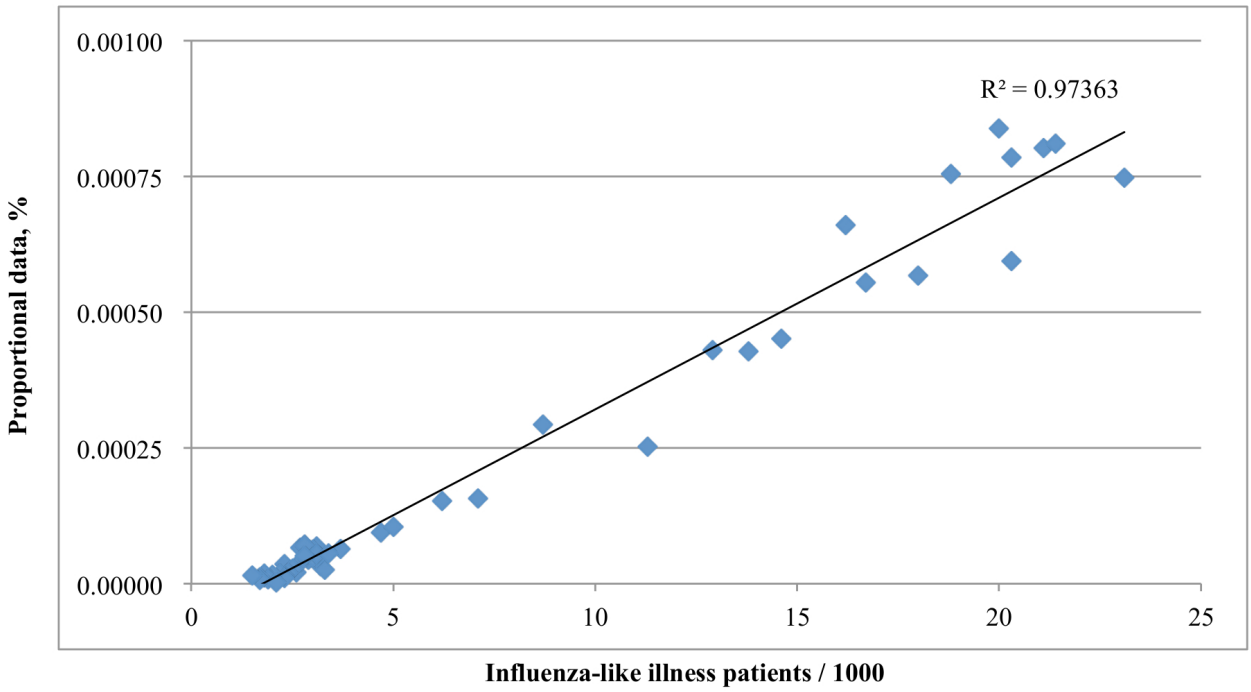


Table 5. Correlation coefficients of combined queries for which the correlation coefficients were ≥.7 and cumulative query methods in set 1.

Cumulative query method from development set 1 (2009/10)	Correlation coefficient		Combined query from development set 1 (2009/10)	Correlation coefficient	
	2009/10	2010/11		2009/10	2010/11
1	.894	.622	New flu (abbr.)[a]	.894	.622
2	.887	.183	Flu + Vaccine	.871	-.062
3	.883	.916	New flu + Cough	.849	.93
4	.861	.933	New flu + Fever	.814	.591
5	.86	.933	Tamiflu + Vaccine	.805	-.062
6	.859	.933	Tamiflu + Symptom	.8	.[b]
7	.849	.943	Flu + Symptom	.799	.815
8	.849	.943	H1N1 + Symptom	.791	.[b]
9	.851	.943	New flu + Sore throat	.738	.504
10	.853	.943	New flu (abbr.)[a] + Vaccine	.713	.[b]
11	.712	.838	New flu + Symptom	.709	.836
12	.728	.841	Tamiflu	.703	.904
13	.728	.841	Tamiflu (English)[c]	.7	.523

[a]abbr.: abbreviation

[b]Correlation cannot be computed because it has a constant value in that period (see Multimedia Appendix 2).

[c]The query was originally submitted in English.

XSL•FO
RenderX

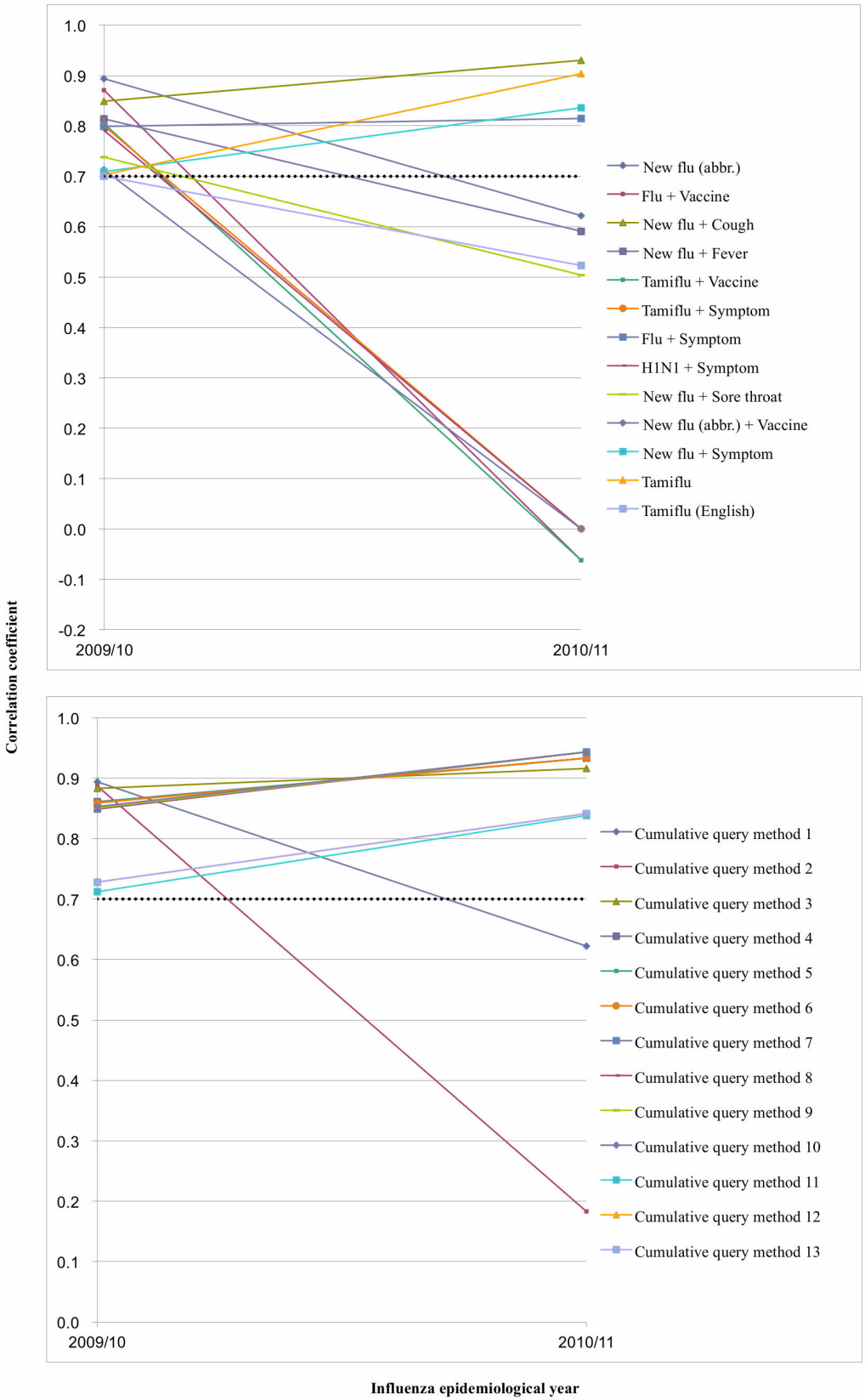
Table 6. Correlation coefficients of combined queries for which the correlation coefficients were ≥.7 and cumulative query methods in set 2.

Cumulative query method from development set 2 (2010/11)	Correlation coefficient		Combined query from development set 2 (2010/11)	Correlation coefficient	
	2010/11	2011/12		2010/11	2011/12
1	.969	.981	Bad cold + Symptom	.969	.981
2	.977	.975	New flu + Treatment	.951	.616
3	.978	.975	New flu + Cough	.93	.291
4	.982	.975	New flu + Sign	.919	.684
5	.97	.987	Tamiflu	.904	.981
6	.968	.986	New influenza + Symptom	.896	.65
7	.969	.986	Bad cold + Treatment	.887	.814
8	.967	.986	Swine flu + Symptom	.877	.005
9	.853	.968	New flu + Symptom	.836	.936
10	.853	.968	Flu + Symptom	.815	.416
11	.854	.965	Influenza + Symptom	.813	.782
12	.854	.965	Influenza (English)[a]	.762	.751
13	.857	.964	New influenza	.748	.503
14	.857	.964	Bird flu + Symptom	.747	.005
15	.86	.78	Bird flu	.709	.136

[a]The query was originally submitted in English.

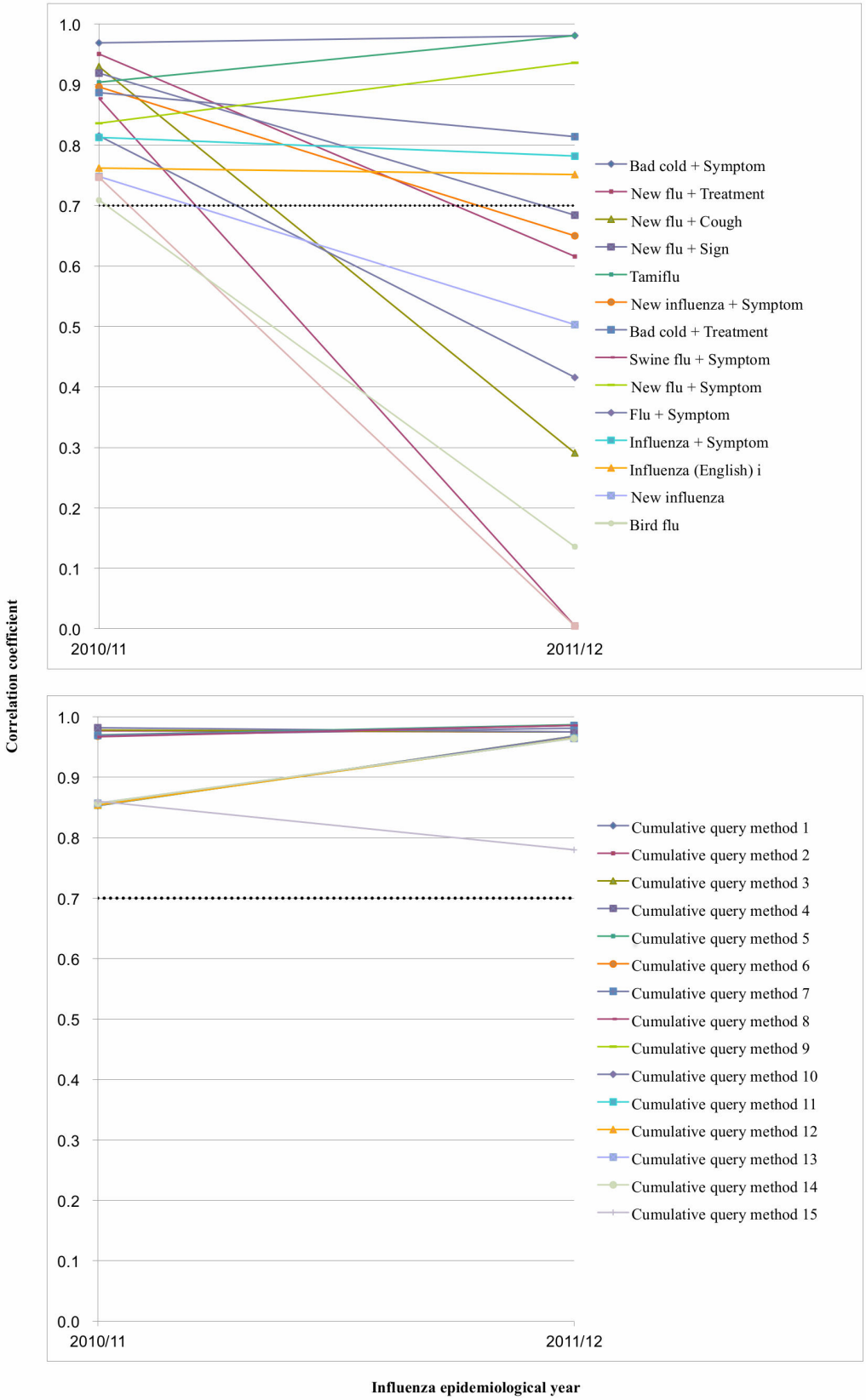
XSL•FO
RenderX

Figure 3. Plot of combined queries for which the correlation coefficients were .7 or higher and cumulative query methods of set 1.



XSL•FO
RenderX

Figure 4. Plot of combined queries for which the correlation coefficients were .7 or higher and cumulative query methods of set 2.



XSL•FO
RenderX

Discussion

Principal Findings

In this study, the cumulative query method showed relatively higher correlation with national influenza surveillance data than combined queries in the development and validation set.

Many people use Internet searches for health information before visiting a doctor [18,20,23,33]. Hence, search query trends can reflect actual disease progression earlier than conventional surveillance. Queries used prior to this study only reflected the authors' opinions [13] or were obtained from databases [13,14,23,37]. To obtain population search queries, we carried out a study survey.

Search queries may vary from country to country. In Korea, "Bad cold (독감)" in Korean has two meanings: one is influenza and the other, a severe common cold. Since the 2009/10 epidemiological season, the Influenza A (H1N1) pandemic period, the media began to use "New flu (신종플루)" in order to distinguish H1N1 influenza and previous influenzas. In 2010, KCDC announced that the official term was "Influenza (인플루엔자)" [36]. But "New flu (신종플루)" and "Bad cold (독감)" are still more popular terms than "Flu (플루)" or "Influenza (인플루엔자)" in Korea (Table 1).

For the 2009/10 epidemiological year (development set 1), 13 combined queries had correlation coefficient r values ≥.7. However, only 4 of these combined queries ("New flu + Cough", "Flu + Symptom", "New flu + Symptom", and "Tamiflu") had correlation coefficient r values ≥.7 in the 2010/11 epidemiological year (validation set 1) (Table 3). But 11 of 13 cumulative query methods had an r value of ≥.7 in validation set 1 (see Table 5 and Figure 3). Among 15 combined queries of development set 2, the number of the combined queries that had correlation coefficient r values of ≥.7 reduced to 6 in validation set 2. But all 15 cumulative query methods had an r value of ≥.7 in validation set 2. We think that the cumulative query method is more robust with time, and this factor is helpful for improving surveillance performance using search queries. Since Internet searching behavior may change over time, this could have affected the performance of Web query-based surveillance model [31]. In this study, 20 out of 210 combined queries had correlation coefficients for all 3 years. And only "Tamiflu" and "New flu + Symptom" showed correlation coefficients r values of ≥.7 for 3 consecutive years (see Figure 1 and Multimedia Appendix 2). Recently, a study using Google Trends for influenza surveillance showed that Google Trends can be used as a complementary source of data [33]; however, its performance is insufficient for use as a model for prediction because its maximum correlation coefficient was .82 for only one query, "Fever", in 2009, and the coefficient decreased to .64 in 2011 [33].

It is difficult to predict the change of search queries in the future. To reduce the effects from changes in search queries, we used a combination of queries and cumulation of combined queries to construct our method. Additionally, the method we wanted to develop was meaningful only when the cumulative query method had a higher correlation coefficient than the highest single combined query. In each validation set, 8 useful cumulative query methods were developed. The useful cumulative query methods in each validation set had a high correlation coefficient (Table 4). In validation set 2, the range of correlation coefficients of the useful cumulative query methods was from .975 to .987. These values are similar to or higher than those reported elsewhere [13,14,23,26,27,29,31]. In Europe, correlation coefficients of .716 to .940 were reported for GFT [27], and coefficients of .82 to .99 were reported in the United States [23,26,31]. In the 2009/10 epidemiological year, called the Influenza A (H1N1) pandemic period, the proportional data of queries were likely to have been different compared to the other epidemiological year. It might affect performance of cumulative query methods in set 1. The performance of the cumulative query method in set 1 was decreased with time (Table 4). It is thought to be related to the changes of queries (see Table 3). For some cumulative query methods, the correlation did not increase by adding queries. The added query did not give extra value in the cumulative query methods 6, 8, and 10 in validation set 1 (see Table 4 and Multimedia Appendix 2). Combined queries 6, 8, and 10 from development set 1 in validation set 1 have a constant value 0 (see Table 3 and Multimedia Appendix 2). The added queries were relatively too small compare to the previous queries in the cumulative query methods 5, 9, 10 in validation sets 1, 3, 4, 7, 8, 10, 12, and 14 in validation set 2 (see Table 4 and Multimedia Appendix 2).

We used proportional data from Daum, a non-dominant local search engine (approximately 25% of the market share) in South Korea [32]. Our cumulative query methods showed a strong correlation with KCDC ILI data. Generally, Google's market share is dominant in countries where GFT is available [27,28]. Our study showed the possibility of developing a surveillance model using a non-dominant local search engine.

Limitations

There are several limitations to this study. The survey of our study is not a representative sample. Because respondents were asked to provide typed queries without mention of the influenza pandemic of 2009/10, recent search queries were more likely to have been included in this study because the survey was conducted recently. This might affect performance of the cumulative query method. Further, the data from the influenza pandemic of 2009/10 might affect the outcome of this study. In this study, we did not combine queries from the same query group. Although important, the performance of using symptoms in the definition of KCDC ILI was not tested. The learning effect from the influenza pandemic of 2009/10, news reports, outbreak briefs, health information from the Internet, and changing search behavior stemming from the diffusion of smartphones might have affected the outcome of this study. We did not determine the extent to which these factors affected the searching behavior. More data for subsequent years are required in order to know the life of the cumulative query method.

Conclusion

We presented a cumulative query method using search engine data. We conducted a survey to obtain population search queries. To reduce the effects from changes in search queries, we used a combination of queries and cumulation of combined queries.

XSL•FO
RenderX

Our method showed high correlation with national influenza surveillance data in South Korea. However, to further our method, additional research is needed.

Acknowledgments

This study was supported by grant 2012-0580 from the Asan Institute for Life Sciences, Seoul, Korea. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Asan Institute for Life Sciences, Seoul, Korea. The technical consultation was supported by Daum Communications. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflicts of Interest

Our study was based on the search engine Daum. This study was partly supported by Daum Communications, the employer of author Maengsoo Yu.

Multimedia Appendix 1

Seasonality of influenza in South Korea.

[PDF File (Adobe PDF File), 47KB - jmir_v16i12e289_app1.pdf]

Multimedia Appendix 2

Full study data (proportional data from Daum are multiplied by 12 squares of 10).

[XLSX File (Microsoft Excel File), 393KB - jmir_v16i12e289_app2.xlsx]

Multimedia Appendix 3

Scaled proportional data based on the two best cumulative query methods.

[PDF File (Adobe PDF File), 143KB - jmir_v16i12e289_app3.pdf]

Multimedia Appendix 4

Scatter plots between the KCDC ILI and other useful cumulative query methods.

[PDF File (Adobe PDF File), 397KB - jmir_v16i12e289_app4.pdf]

Multimedia Appendix 5

Cumulative query method for influenza virologic data.

[PDF File (Adobe PDF File), 54KB - jmir_v16i12e289_app5.pdf]

References

1. Triple S Project. Assessment of syndromic surveillance in Europe. Lancet 2011 Nov 26;378(9806):1833-1834. [doi: 10.1016/S0140-6736(11)60834-9] [Medline: 22118433]
2. Hirshon JM. The rationale for developing public health surveillance systems based on emergency department data. Acad Emerg Med 2000 Dec;7(12):1428-1432. [Medline: 11099436]
3. Henning KJ. What is syndromic surveillance? MMWR Morb Mortal Wkly Rep 2004 Sep 24;53 Suppl:5-11 [FREE Full text] [Medline: 15714620]
4. Ferguson NM, Cummings DA, Cauchemez S, Fraser C, Riley S, Meeyai A, et al. Strategies for containing an emerging influenza pandemic in Southeast Asia. Nature 2005 Sep 8;437(7056):209-214. [doi: 10.1038/nature04017] [Medline: 16079797]
5. Longini IM, Nizam A, Xu S, Ungchusak K, Hanshaoworakul W, Cummings DA, et al. Containing pandemic influenza at the source. Science 2005 Aug 12;309(5737):1083-1087 [FREE Full text] [doi: 10.1126/science.1115717] [Medline: 16079251]
6. Cheng CK, Cowling BJ, Lau EH, Ho LM, Leung GM, Ip DK. Electronic school absenteeism monitoring and influenza surveillance, Hong Kong. Emerg Infect Dis 2012 May;18(5):885-887 [FREE Full text] [doi: 10.3201/eid1805.111796] [Medline: 22516519]

XSL•FO
RenderX

7. Egger JR, Hoen AG, Brownstein JS, Buckeridge DL, Olson DR, Konty KJ. Usefulness of school absenteeism data for predicting influenza outbreaks, United States. Emerg Infect Dis 2012 Aug;18(8):1375-1377 [FREE Full text] [doi: 10.3201/eid1808.111538] [Medline: 22840354]
8. Galante M, Garin O, Sicuri E, Cots F, García-Altés A, Ferrer M, et al. Health services utilization, work absenteeism and costs of pandemic influenza A (H1N1) 2009 in Spain: a multicenter-longitudinal study. PLoS One 2012;7(2):e31696 [FREE Full text] [doi: 10.1371/journal.pone.0031696] [Medline: 22348122]
9. Ohkusa Y, Shigematsu M, Taniguchi K, Okabe N. Experimental surveillance using data on sales of over-the-counter medications--Japan, November 2003-April 2004. MMWR Morb Mortal Wkly Rep 2005 Aug 26;54 Suppl:47-52 [FREE Full text] [Medline: 16177693]
10. Patwardhan A, Bilkovski R. Comparison: Flu prescription sales data from a retail pharmacy in the US with Google Flu trends and US ILINet (CDC) data as flu activity indicator. PLoS One 2012;7(8):e43611 [FREE Full text] [doi: 10.1371/journal.pone.0043611] [Medline: 22952719]
11. Vergu E, Grais RF, Sarter H, Fagot JP, Lambert B, Valleron AJ, et al. Medication sales and syndromic surveillance, France. Emerg Infect Dis 2006 Mar;12(3):416-421 [FREE Full text] [doi: 10.3201/eid1205.050573] [Medline: 16704778]
12. Hill S, Mao J, Ungar L, Hennessy S, Leonard CE, Holmes J. Natural supplements for H1N1 influenza: retrospective observational infodemiology study of information and search activity on the Internet. J Med Internet Res 2011;13(2):e36 [FREE Full text] [doi: 10.2196/jmir.1722] [Medline: 21558062]
13. Eysenbach G. Infodemiology: tracking flu-related searches on the web for syndromic surveillance. AMIA Annu Symp Proc 2006:244-248 [FREE Full text] [Medline: 17238340]
14. Hulth A, Rydevik G. Web query-based surveillance in Sweden during the influenza A(H1N1)2009 pandemic, April 2009 to February 2010. Euro Surveill 2011;16(18):- [FREE Full text] [Medline: 21586265]
15. Yang AC, Huang NE, Peng CK, Tsai SJ. Do seasons have an influence on the incidence of depression? The use of an internet search engine query data as a proxy of human affect. PLoS One 2010;5(10):e13728 [FREE Full text] [doi: 10.1371/journal.pone.0013728] [Medline: 21060851]
16. 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]
17. Chretien JP, George D, Shaman J, Chitale RA, McKenzie FE. Influenza forecasting in human populations: a scoping review. PLoS One 2014;9(4):e94130 [FREE Full text] [doi: 10.1371/journal.pone.0094130] [Medline: 24714027]
18. 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]
19. Zheluk A, Gillespie JA, Quinn C. Searching for truth: internet search patterns as a method of investigating online responses to a Russian illicit drug policy debate. J Med Internet Res 2012;14(6):e165 [FREE Full text] [doi: 10.2196/jmir.2270] [Medline: 23238600]
20. Bernardo TM, Rajic A, Young I, Robiadek K, Pham MT, Funk JA. Scoping review on search queries and social media for disease surveillance: a chronology of innovation. J Med Internet Res 2013;15(7):e147 [FREE Full text] [doi: 10.2196/jmir.2740] [Medline: 23896182]
21. Liang B, Scammon DL. Incidence of online health information search: a useful proxy for public health risk perception. J Med Internet Res 2013;15(6):e114 [FREE Full text] [doi: 10.2196/jmir.2401] [Medline: 23773974]
22. Yom-Tov E, Gabrilovich E. Postmarket drug surveillance without trial costs: discovery of adverse drug reactions through large-scale analysis of web search queries. J Med Internet Res 2013;15(6):e124 [FREE Full text] [doi: 10.2196/jmir.2614] [Medline: 23778053]
23. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. Nature 2009 Feb 19;457(7232):1012-1014. [doi: 10.1038/nature07634] [Medline: 19020500]
24. Eurosurveillance editorial team. Google Flu Trends includes 14 European countries. Euro Surveill 2009;14(40):- [FREE Full text] [Medline: 19822118]
25. Malik MT, Gumel A, Thompson LH, Strome T, Mahmud SM. "Google flu trends" and emergency department triage data predicted the 2009 pandemic H1N1 waves in Manitoba. Can J Public Health 2011;102(4):294-297. [Medline: 21913587]
26. Ortiz JR, Zhou H, Shay DK, Neuzil KM, Fowlkes AL, Goss CH. Monitoring influenza activity in the United States: a comparison of traditional surveillance systems with Google Flu Trends. PLoS One 2011;6(4):e18687 [FREE Full text] [doi: 10.1371/journal.pone.0018687] [Medline: 21556151]
27. Valdivia A, Lopez-Alcalde J, Vicente M, Pichiule M, Ruiz M, Ordobas M. Monitoring influenza activity in Europe with Google Flu Trends: comparison with the findings of sentinel physician networks - results for 2009-10. Euro Surveill 2010;15(29):- [FREE Full text] [Medline: 20667303]
28. Wilson N, Mason K, Tobias M, Peacey M, Huang QS, Baker M. Interpreting Google flu trends data for pandemic H1N1 influenza: the New Zealand experience. Euro Surveill 2009;14(44):- [FREE Full text] [Medline: 19941777]
29. Timpka T, Spreco A, Dahlström Ö, Eriksson O, Gursky E, Ekberg J, et al. Performance of eHealth data sources in local influenza surveillance: a 5-year open cohort study. J Med Internet Res 2014;16(4):e116 [FREE Full text] [doi: 10.2196/jmir.3099] [Medline: 24776527]

XSL•FO
RenderX

30. Pervaiz F, Pervaiz M, Abdur Rehman N, Saif U. FluBreaks: early epidemic detection from Google flu trends. J Med Internet Res 2012;14(5):e125 [FREE Full text] [doi: 10.2196/jmir.2102] [Medline: 23037553]
31. Cook S, Conrad C, Fowlkes AL, Mohebbi MH. Assessing Google flu trends performance in the United States during the 2009 influenza virus A (H1N1) pandemic. PLoS One 2011;6(8):e23610 [FREE Full text] [doi: 10.1371/journal.pone.0023610] [Medline: 21886802]
32. Market share of search engine in South Korea. URL: http://www.webWebcitation.org/6QrnPuRwJ [accessed 2014-12-02] [WebCite Cache ID 6UVV1DSZJ]
33. Kang M, Zhong H, He J, Rutherford S, Yang F. Using Google Trends for influenza surveillance in South China. PLoS One 2013;8(1):e55205 [FREE Full text] [doi: 10.1371/journal.pone.0055205] [Medline: 23372837]
34. Cho S, Sohn CH, Jo MW, Shin SY, Lee JH, Ryoo SM, et al. Correlation between national influenza surveillance data and google trends in South Korea. PLoS One 2013;8(12):e81422 [FREE Full text] [doi: 10.1371/journal.pone.0081422] [Medline: 24339927]
35. Search engine Daum. URL: http://www.webWebcitation.org/6RvfPYTwI [accessed 2014-12-01] [WebCite Cache ID 6UVV8V6Dr]
36. Korea Centers for Disease Control and Prevention. URL: http://www.webWebcitation.org/6QrlqwNxO [accessed 2014-12-01] [WebCite Cache ID 6UVVAvSvZ]
37. Polgreen PM, Chen Y, Pennock DM, Nelson FD. Using internet searches for influenza surveillance. Clin Infect Dis 2008 Dec 1;47(11):1443-1448 [FREE Full text] [doi: 10.1086/593098] [Medline: 18954267]

Abbreviations

GFT: Google Flu Trends
ILI: influenza-like illness
KCDC: Korea Centers for Disease Prevention and Control
PCR: polymerase chain reaction

Edited by G Eysenbach; submitted 07.07.14; peer-reviewed by M Kang, E Lau; comments to author 28.07.14; revised version received 25.08.14; accepted 21.11.14; published 16.12.14.

Please cite as:
Seo DW, Jo MW, Sohn CH, Shin SY, Lee J, Yu M, Kim WY, Lim KS, Lee SI
Cumulative Query Method for Influenza Surveillance Using Search Engine Data
J Med Internet Res 2014;16(12):e289
URL: http://www.jmir.org/2014/12/e289/
doi:10.2196/jmir.3680
PMID:25517353

©Dong-Woo Seo, Min-Woo Jo, Chang Hwan Sohn, Soo-Yong Shin, JaeHo Lee, Maengsoo Yu, Won Young Kim, Kyoung Soo Lim, Sang-Il Lee. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 16.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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 Home Telemonitoring to Support Multidisciplinary Care of Heart Failure Patients in Finland: Randomized Controlled Trial

Anna-Leena Vuorinen[1], MSc; Juha Leppänen[2], MSc; Hannu Kaijanranta[2], MSc; Minna Kulju[1], MSc; Tiina Heliö[3], MD, PhD; Mark van Gils[1], PhD; Jaakko Lähteenmäki[2], Tech. Lic

[1]VTT Technical Research Centre of Finland, Tampere, Finland

[2]VTT Technical Research Centre of Finland, Espoo, Finland

[3]Helsinki University Central Hospital, Heart and Lung Center, Helsinki, Finland

Corresponding Author:
Anna-Leena Vuorinen, MSc
VTT Technical Research Centre of Finland
PO Box 1300
Tampere, FIN-33101
Finland
Phone: 358 40 848 5966
Fax: 358 20 722 7001
Email: anna-leena.vuorinen@vtt.fi

Abstract

Background: Heart failure (HF) patients suffer from frequent and repeated hospitalizations, causing a substantial economic burden on society. Hospitalizations can be reduced considerably by better compliance with self-care. Home telemonitoring has the potential to boost patients' compliance with self-care, although the results are still contradictory.

Objective: A randomized controlled trial was conducted in order to study whether the multidisciplinary care of heart failure patients promoted with telemonitoring leads to decreased HF-related hospitalization.

Methods: HF patients were eligible whose left ventricular ejection fraction was lower than 35%, NYHA functional class ≥2, and who needed regular follow-up. Patients in the telemonitoring group (n=47) measured their body weight, blood pressure, and pulse and answered symptom-related questions on a weekly basis, reporting their values to the heart failure nurse using a mobile phone app. The heart failure nurse followed the status of patients weekly and if necessary contacted the patient. The primary outcome was the number of HF-related hospital days. Control patients (n=47) received multidisciplinary treatment according to standard practices. Patients' clinical status, use of health care resources, adherence, and user experience from the patients' and the health care professionals' perspective were studied.

Results: Adherence, calculated as a proportion of weekly submitted self-measurements, was close to 90%. No difference was found in the number of HF-related hospital days (incidence rate ratio [IRR]=0.812, P=.351), which was the primary outcome. The intervention group used more health care resources: they paid an increased number of visits to the nurse (IRR=1.73, P<.001), spent more time at the nurse reception (mean difference of 48.7 minutes, P<.001), and there was a greater number of telephone contacts between the nurse and intervention patients (IRR=3.82, P<.001 for nurse-induced contacts and IRR=1.63, P=.049 for patient-induced contacts). There were no statistically significant differences in patients' clinical health status or in their self-care behavior. The technology received excellent feedback from the patient and professional side with a high adherence rate throughout the study.

Conclusions: Home telemonitoring did not reduce the number of patients' HF-related hospital days and did not improve the patients' clinical condition. Patients in the telemonitoring group contacted the Cardiology Outpatient Clinic more frequently, and on this way increased the use of health care resources.

Trial Registration: Clinicaltrials.gov NCT01759368; http://clinicaltrials.gov/show/NCT01759368 (Archived by WebCite at http://www.webcitation.org/6UFxiCk8Z).

(J Med Internet Res 2014;16(12):e282) doi:10.2196/jmir.3651

KEYWORDS

heart failure; telemonitoring; hospitalization; user experience; clinical outcomes; EHFSBS; health care resources

XSL•FO
RenderX

Introduction

Heart failure (HF) is a serious and costly disease associated with poor quality of life [1], a wide range of comorbidities [2], and a high rate of hospitalization [3]. Nearly 25% of patients are readmitted within 30 days [4], and by 6 months, the proportion increases to 50% [5]. Hospitalizations cause a heavy economic burden since they are responsible for 60-70% of the total costs of HF care [6]. Moreover, the 1-year mortality of HF patients is 30% [3], and the 5-year survival rate is poorer than in most cancers [7].

A multidisciplinary care approach to heart failure is incorporated with European and American guidelines. The multidisciplinary care model includes specially trained HF nurses, the education of patients (and caregivers) regarding precipitating factors and the need for compliance with medication and diet, follow-up monitoring by trained staff, and access to specialized HF clinics [8]. Non-compliance with medication and other lifestyle recommendations is a major problem among HF patients resulting in worsening symptoms that can lead to readmission [9]. Hospitalizations may be preventable by up to 50% mainly by improving compliance with self-care [10].

Care-delivery models that incorporate telemonitoring as a part of HF patients' care have the potential to boost patients' compliance with self-care while at the same time bringing health care services closer to them. Meta-analyses from the years 2009-2011 link telemonitoring with improved survival, decreased hospitalizations, and improved quality of life [11-13]. However, since these meta-analyses were carried out there have been two large randomized controlled trials that have failed to show evidence in favor of telemonitoring in terms of reducing hospitalizations and death [14,15]. Similar findings have been reported in earlier studies [16,17] and more recently in smaller studies [18-20], except in the TEN-HMS trial [16], in which mortality was found to be lower in the telemonitoring group compared to usual care. Furthermore, results from the recent Whole Systems Demonstrator (WSD) study, a multisite trial involving 3230 chronically ill patients shows contradictive evidence. Telehealth was found to reduce mortality and emergency admission rates in the secondary care [21] but failed to improve quality of life or psychological outcomes [22], nor was it cost-effective [23]. Among patients with social care needs in WSD study, telecare did not alter the use of health and social care service or mortality [24]. To summarize, the literature shows conflicting evidence on the effectiveness of telemonitoring dependent on the target population and study environment and the implementation and structure of the intervention itself.

The current literature does not cover the evaluation of telemonitoring as a part of multidisciplinary care. The objective of this study was to investigate whether the multidisciplinary care of HF patients could be improved with telemonitoring at the Cardiology Outpatient Clinic of Helsinki University Central Hospital (HUCH), primarily in terms of reducing HF-related hospitalizations. We hypothesized that telemonitoring improves patients' adherence to self-care—something that will be realized as decreased hospitalizations.

Methods

Study Design

Heart at Home was a two-arm randomized controlled trial conducted at the Cardiology Outpatient Clinic of HUCH in 2010-2012 (NCT01759368). The study protocol was approved by the Ethics Committee of the Hospital District of Helsinki and Uusimaa. All the patients provided a written informed consent before they were randomized. (See Multimedia Appendix 1 for the CONSORT EHEALTH checklist [25]).

Matched pair design was used in the randomization. The eligible patients, who were similar in left ventricular ejection fraction, NYHA classification, age, and gender, respectively, were matched in pairs. One was randomized to the control group and the other to the intervention group.

The study was divided into two parts. The first 30 intervention patients and 29 control patients started stepwise from November 2010 to February 2011. After the first 59 patients had finished their follow-up, the second group (17 intervention patients and 18 control patients) started in May to August 2011. The nominal follow-up time was 6 months. The study was completed in February 2012.

Participants

Patients suffering from chronic heart failure were recruited to the study. The inclusion criteria were (1) diagnosis of systolic heart failure, (2) age of 18-90 years, (3) NYHA class ≥2 (an interview-based classification by the New York Heart Association concerning limitations to physical activity), (4) left ventricular ejection fraction ≤35% as measured during hospital visits, (5) need for a regular check-up visit, and (6) time from the last visit of less than 6 months. Patients were not eligible if they had a planned major medical operation, had severe comorbidity such as cancer, had participated in another clinical trial during the last 3 months, or were suspected of poor compliance. The assessment of compliance was based on patient's technical skills, such as ability to use a mobile phone.

The electronic patient database of HUCH was used for the initial screening of patients with chronic heart failure so as to further assess their eligibility. Eligible patients were informed about the study and were asked whether they were willing to participate (and their formal consent was obtained) when they came for their normal follow-up visit. For willing patients, anthropometric and laboratory measurements were taken and the patients completed the study questionnaires. For each patient, the medication was checked and optimized. The same procedure was repeated at the end-point visit.

Usual Care

At the Cardiology Outpatient Clinic of HUCH, there are about 600 HF patients, of whom 150-200 patients have serious heart failure that requires regular follow-up visits. A multidisciplinary care approach including patient guidance and support for self-care has been adopted at the clinic. In the care of these HF patients, the cardiac team plays a central role in monitoring and interpreting patient symptoms, optimizing medication, and providing education. The cardiac team consists of 2 physicians,

XSL•FO
RenderX

a specialized heart failure nurse, and a physiotherapist who helps after a hospitalization period. As part of the care process, patients capable of carrying out self-care were identified and encouraged to regularly measure their blood pressure, heart rate, and weight at home. The information exchange between HF patients and care personnel took place during patients' visits to the clinic and by telephone. Systematic collection and exploitation of the self-measurement data was difficult since it depended on the patient's own activity. Often a patient had not monitored their health parameters as agreed or had forgotten to bring along the measurement notes. The heart failure nurse contacted patients by telephone if agreed in the care plan to motivate and remind them to comply with the self-care plan.

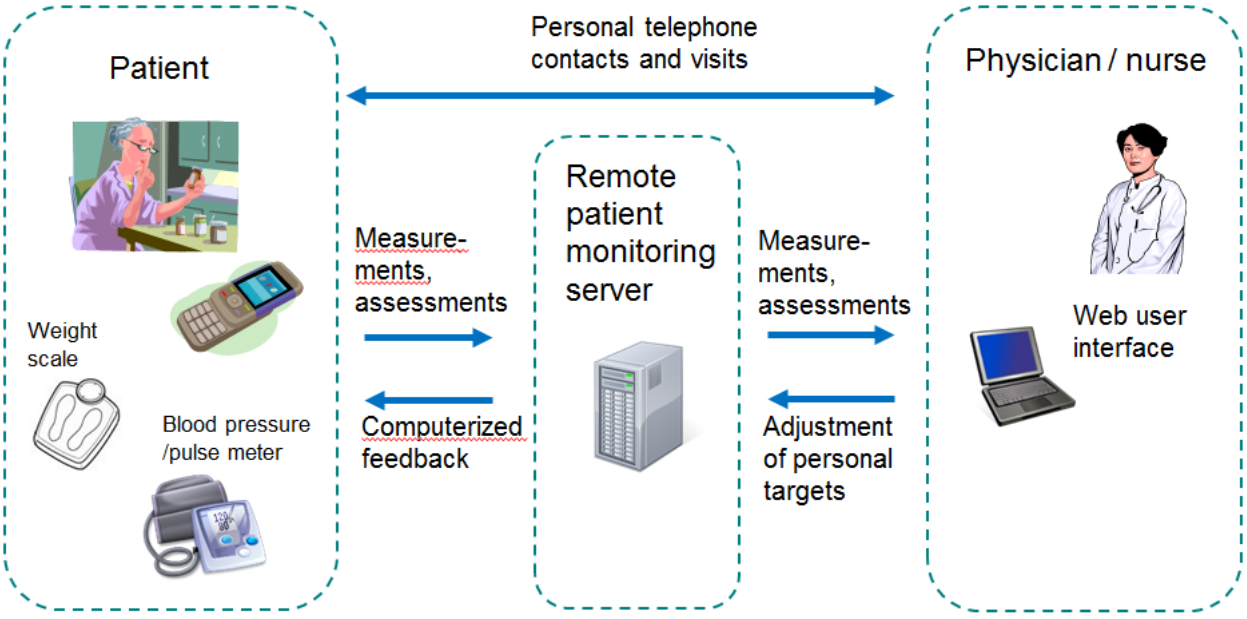
Intervention: A Telemonitoring-Assisted Self-Care Model

For patients in the intervention arm, a new care process was introduced in which a patient regularly reported their most important health parameters to the nurse using a mobile phone app. At the beginning of the study, the patients were given a home-care package including a weight scale, a blood pressure meter, a mobile phone, and self-care instructions. The patients were advised to carry out and report the measurements together with the assessment of symptoms once a week.

A pre-installed software app on the mobile phone supported the uploading of measurements and the self-assessment of symptoms. In the development of the mobile app, particular care was paid to the simplicity of the user interface and its ease of use, since most of the patients were elderly. The measurements taken at home to be uploaded were diastolic and systolic blood pressure, pulse, body weight, and an assessment of symptoms. The symptom assessment concerned the patient's feelings of dizziness, dyspnea, palpitation, weakness, and edema. Patients were also asked to evaluate their overall condition, that is, whether their condition had deteriorated, improved, or remained unchanged. In the context of each submission of information, the patient received automatic machine-based feedback of whether the reported parameter was within their personal targets set by the nurse. The overall architecture used in the self-care process and screenshots of the software app are depicted in Figures 1 and 2. The system was developed by VTT Technical Research Centre of Finland.

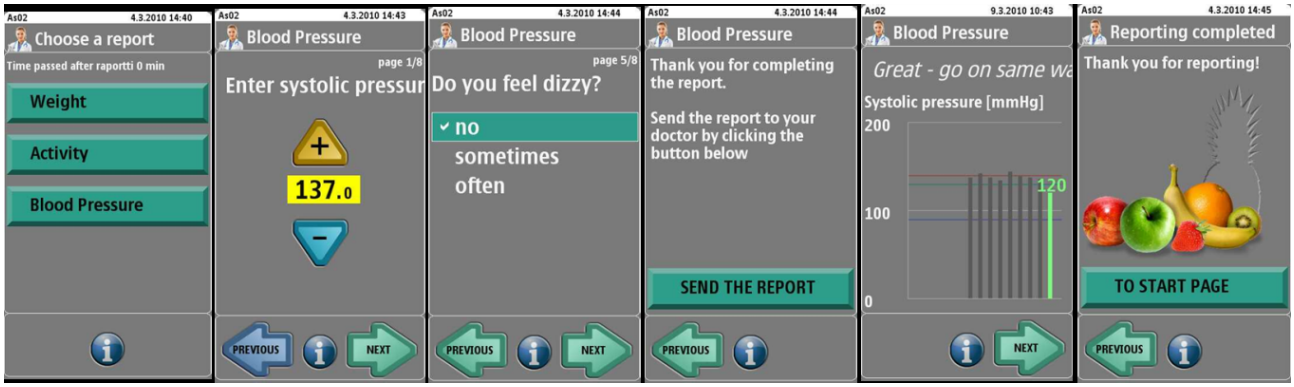
The measurements were stored on the secured remote patient monitoring server. The cardiac team was able to access the data with a browser-based user interface. The nurse followed the patient's status and the data once a week or more frequently if necessary. In the beginning of the study, the nurse contacted the patient every time the measurement was beyond the target levels or if the patient reported any of the symptoms. Later, the contacts were more dependent on the patient's measurement history. If the latest measurement markedly differed from previous measurements, the nurse called the patient. The nurse could invite the patient for a check-up visit if still necessary after the phone call. If a patient did not comply with the weekly reporting plan, the nurse contacted the patient and encouraged him or her to continue with the monitoring.

Figure 1. Overall architecture for remote patient monitoring.



XSL•FO
RenderX

Figure 2. Screenshots of the reporting process with the mobile app.



Outcome Measures

Primary Outcome

The primary outcome was the number of HF-related hospital days during the follow-up. The data were obtained from the electronic health record system of HUCH.

Secondary Outcomes

Secondary outcomes included clinical outcomes, use of health care resources, and user experience. The following variables were analyzed in order to assess clinical effectiveness: death from any cause, heart transplant operation or listing for transplant operation, left ventricular ejection fraction (LVEF, %) measured by echocardiography, plasma concentration of N-terminal of the prohormone brain natriuretic peptide (NT-proBNP, ng/l), creatinine (µmol/l), sodium (mmol/l), and potassium (mmol/l). For the plasma concentrations of sodium, potassium, and creatinine, there is no unambiguous interpretation of the direction of change, but the value should be within the reference range. The reference ranges used at HUCH are sodium 137-145 mmol/l, potassium 3.3-4.9 mmol/l, creatinine among women 50-90 µmol/l and among men 60-100µmol/l. Sodium, potassium, and creatinine were dichotomized indicating whether the observed value was within the reference range.

Self-care behavior was measured using the European Heart Failure Self-Care Behaviour Scale (EHFSBS). EHFSBS is a 12-item self-administered questionnaire specifically designed and tested for HF patients including statements on self-care behavior essential in the care of HF. The statements are scored from 1-5; the lower the score, the better the performance in self-care. The summary score was analyzed, and medication changes were recorded to examine how the telemonitoring intervention affected activity in medication regimen. The nurse collected information regarding changes in patients' medication regimen throughout the study. Changes related to medication optimization during the baseline visit were excluded. Changes made to patients' medication were classified into three categories: increase of medication (a new drug or increase in dosage), decrease of medication (termination of a certain drug or decrease in dosage), and self-imposed medication termination (patient had stopped taking medicine without physician's confirmation). The medications were classified as diuretics, ACE-I, or beta-blockers.

In terms of the use of health care resources, outpatient visits were analyzed: the number of (1) unplanned visits to the Cardiology Outpatient Clinic (nurse or physician), (2) visits to the emergency department, (3) visits to and time spent with the nurse, (4) visits to and time spent with the physicians, and (5) telephone contacts made by the patient and by the nurse. The baseline visits and the end-point visits were included in the calculations. The data were retrieved from the electronic health records and by asking the patient.

Patients' acceptance and experience towards home telemonitoring were evaluated using a questionnaire delivered to patients in the telemonitoring group at the end-point visit. The questionnaire included statements about their experiences with the usability of the mobile phone app, as well as their satisfaction with using the app and the benefits of the telemonitoring-assisted care model. In addition, an in-depth interview was conducted with the nurse responsible in order to assess the user experience from a professional perspective.

Power Calculations

The study was designed to have a power of 90%, an alpha level of .05, and an effect size of 0.5 determined as the expected difference of 3 HF-related hospital days between the study groups (SD 6). A *t* test was used as a calculation framework. With these parameters, we calculated that 44 patients per treatment arm needed to be recruited.

Statistical Analysis

The intention-to-treat principle was applied in statistical analyses. There was one dropout in the intervention group. The patient withdrew from the study shortly after the beginning, and no end-point measurements were available. The patient was excluded in the end-point analyses. All analyses except zero inflated Poisson (ZIP) were carried out using SPSS version 19. ZIP regression models were conducted using R version 2.15.1.

Outcome variables that express counts (eg, HF-related hospital days, visits to the nurse, visits to the physician, number of phone calls, unplanned visits to the clinic) were presented using the mean and a percentage of zero counts. Poisson regression and ZIP regression models were used in order to analyze the difference between the study groups. The Vuong test [26] was used to assess the superiority between Poisson regression and ZIP for each variable. Finally, ZIP regression was used in the analysis of the following variables: number of HF-related hospital days, number of unplanned visits to the clinic, and

XSL•FO
RenderX

telephone contacts initiated by the patient. In all models, the patient's individual study duration (in days) was set as an offset variable, and the control group was used as a reference group. The incidence rate ratio (IRR) and its 95% confidence interval (Cl) were reported.

Repeated contiguous variables were analyzed within and between the study groups. The paired *t* test or Wilcoxon matched-pair signed-rank test in the case of non-normality was used for the analyses of within-group changes. Non-normality was confirmed by the Kolmogorov-Smirnov test. Analysis of covariance was used to investigate differences between the control and the intervention groups with adjustment for baseline values. The 95% Cl and *P* value for the between-group difference were reported.

Results

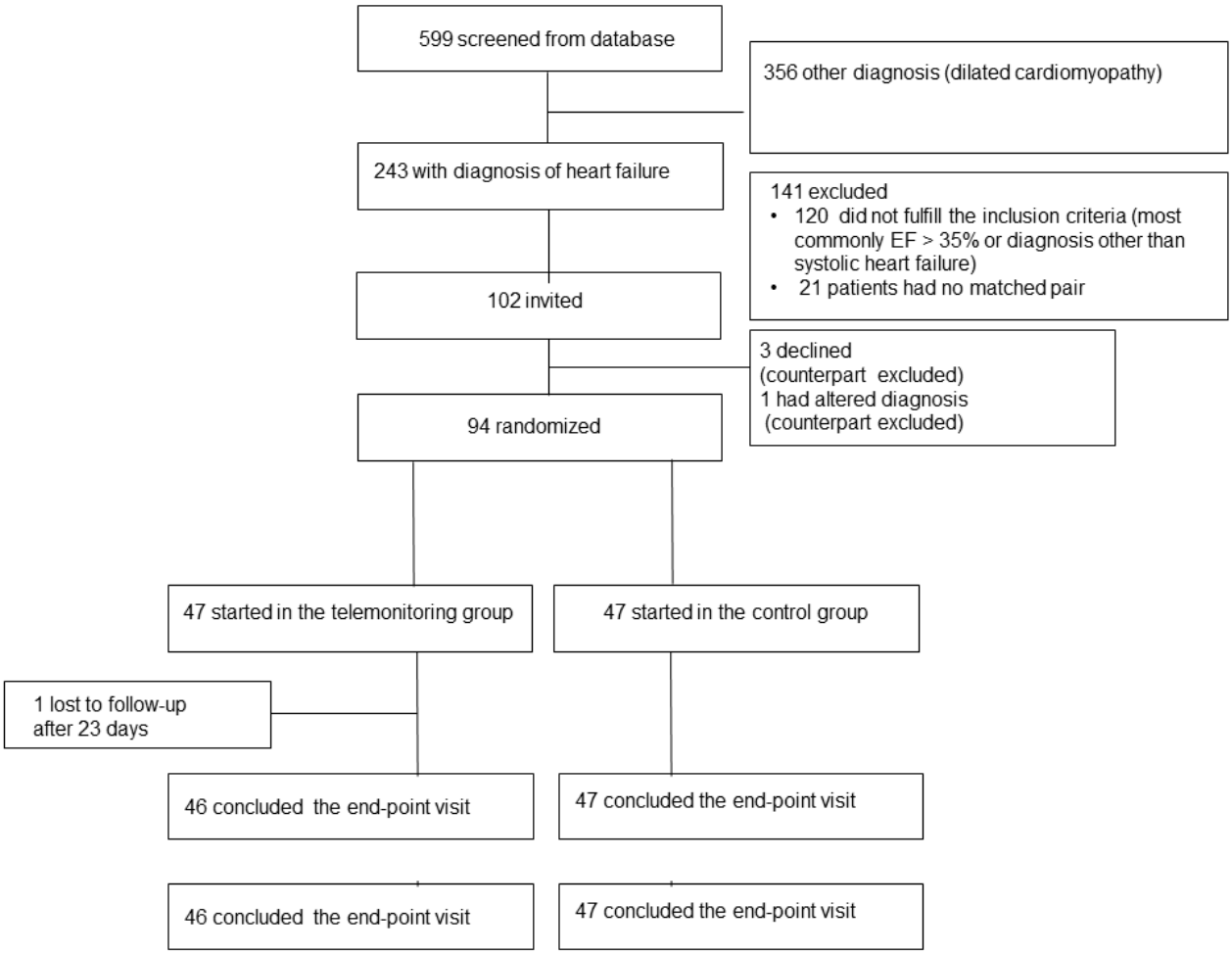
Patient Flow

Figure 3 shows the progress of the study. Altogether, 599 patients were screened from the database, of whom 243 were diagnosed with systolic heart failure. Of these, 123 patients fulfilled the inclusion criteria. Eligible patients who were similar in their left ventricular ejection fraction, NYHA classification, age, and gender were matched; 51 matched pairs were identified. The 102 patients were invited for a baseline visit where baseline measurements were taken and information considering the study was given. Of these, 3 patients declined to participate and another patient had a changed diagnosis. Respectively, their matched counterparts were excluded from the study. Finally, 94 patients were randomized. One from each pair was randomly assigned to receive the usual care, and the other was assigned to the telemonitoring group. There was one dropout in the telemonitoring group. The patient withdrew from the study after 23 days. The patient felt that monitoring his condition made him anxious as it reminded him constantly of the disease.

Baseline Characteristics

Table 1 displays the baseline characteristics of the study subjects in both the control group and the telemonitoring group.

Figure 3. Screening, randomization, and follow-up of patients.



XSL•FO
RenderX

Table 1. Baseline characteristics of the patients: mean (standard deviation) or frequency (percentile).

	Control group (n=47)	Telemonitoring group (n=47)
Sex (male), n (%)	39 (83)	39 (83)
Age in years, mean (SD)	57.9 (11.9)	58.3 (11.6)
BMI (kg/m^2), mean (SD)	27.9 (4.7)	28.4 (6.0)
Systolic blood pressure (mmHg), mean (SD)	116 (16)	112 (13)
Diastolic blood pressure (mmHg), mean (SD)	72 (10)	71 (10)
Heart rate (bpm), mean (SD)	70 (12)	69 (11)
Left ventricular ejection fraction (% units), mean (SD)	28.6 (5.0)	27.3 (4.9)
NYHA, frequency (%)		
Slight limitations in physical activity (Class II)	17 (36)	19 (40)
Marked limitation in physical activity (Class III)	28 (60)	27 (58)
Severe limitations in physical activity (Class IV)	2 (4)	1 (2)
Comorbidities, frequency (%) [a]		
Diabetes	2 (4)	2 (4)
Hypertension	6 (13)	8 (17)
Atrial fibrillation	6 (13)	14 (30)
Asthma/COPD	5 (11)	2 (4)
Renal insufficiency	4 (9)	0 (2)
No comorbidities	9 (19)	12 (26)
Smoking (number of non-smokers), n (%) [a]	42 (89)	35 (76)

[a]Data missing from one patient in the telemonitoring group.

Primary Outcome

On average there were 1.4 (SD 3.5) HF-related hospital days in the control group and 0.7 (SD 2.4) HF-related hospital days in the telemonitoring group. Of the control patients, 72% (34/47), and of the telemonitoring patients, 83% (38/46) had no hospital days during the 6-month follow-up. The difference between the study groups was not statistically significant (IRR=0.812, 95% Cl 0.525-1.256, *P*=.351).

Secondary Outcomes

Contrary to expectations, none of the subjects died, underwent a heart transplant operation, or were listed for a transplant operation. In both study groups, two patients had an emergency episode.

Table 2 shows clinical outcomes at baseline and post intervention. There were no statistically significant differences between the study groups in either of the clinical variables. However, in both study groups, there were significant within-group changes: an increase in LVEF (4.2%, *P*=.001 for control group and 5.0%, *P*=.003 for telemonitoring group) and in EHFSBS (-3.8 points, *P*<.001 in the control group and -5.0, *P*<.001 in the telemonitoring group) and a decrease in NT-proBNP levels in the telemonitoring group (-198ng/l, *P*=.01).

XSL•FO
RenderX

Table 2. Clinical outcomes at the beginning and the end of the study and the within-group change with the corresponding confidence intervals (*P* value refers to the significance level for the between-group difference).

	Control group (n=46)			Telemonitoring group (n=47)				
	Begin	End	Change (95% Cl)	Begin	End	Change (95% Cl)	Effect size (95% Cl)	*P* value
NT-proBNP (ng/l), median (interquartile range)[a]	1338 (474-2974)	731 (368-2408)	-50 (-831 to 260)	2347 (998-3568)	1014 (609-2631)	-198 (-1921 to 170)	Median difference -148	.256
LVEF (%), mean (SD)	28.6 (5.0)	32.8 (8.2)	4.2 (1.8-6.5)	27.3 (4.9)	32.4 (9.8)	5.0 (1.8-8.1)	β=-.04 (-3.658 to 3.743)	.982
EHFSBS score, mean (SD)	27.9 (6.5)	24.1 (8.3)	-3.8 (-5.4 to -2.1)	27.6 (6.8)	22.6 (6.9)	-5.0 (-7.1 to -3.0)	β=-1.320 (-3.842 to 1.184)	.298
Serum creatinine (μmol/L), median (interquartile range; proportion[b])[a]	88.0 (78.0-121.0; 0.60)	92.0 (84.0-115.0; 0.57)	2.4 (-2.5 to 6.5)	92.5 (83.8-105.3; 0.67)	95.0 (83.5-112.3; 0.57)	3.5 (-1.0 to 9.0)	Median difference 0.9	.435
Serum potassium (mmol/L), mean (SD; proportion[b])	4.2 (0.3; 0.98)	4.2 (0.4; 0.98)	-0.06 (-0.18 to 0.05)	4.3 (0.4; 0.94)	4.2 (0.4; 1.0)	-0.1 (-0.3 to 0.01)	β=.008 (-0.141 to 0.156)	.918
Serum sodium (mmol/L), mean (SD; proportion[b])	140.2 (3.0; 0.96)	140.4 (3.4; 0.85)	0.2 (-0.7 to 1.1)	139.6 (3.0; 0.89)	140.7 (2.5; 0.96)	1.0 (0.2-1.8)	β=-.535 (-1.592 to 0.523)	.318

[a]Non-parametric test.

[b]Proportion of patients whose values were within reference interval.

Changes in medication regimen are presented in Table 3. Significantly more medication changes were done to the patients in the telemonitoring group (*P*=.042 for medication increase and *P*=.026 for medication decrease). All decreases in medication were done to telemonitoring patients, and the decreases were applied to diuretics. The increases in medication in the telemonitoring group involved the following types of medication: five increases in angiotensin converting enzyme inhibitor (ACE-I) therapy, three increases in beta-blockers, and two increases in diuretics. In the control group, the two increases were applied to diuretics.

Table 3. Categorized medication adjustments and the number of patients to whom the adjustments were applied.

Adjustments	Control group (n=47)	Telemonitoring group (n=46)	*P* value
Increase in medication, n (%)	2 (4)	8 (17)	.042
Decrease in medication, n (%)	0 (0)	5 (11)	.026
Self-imposed medication termination, n (%)	3 (6)	2 (4)	1.0

Table 4 shows the use of health care resources. The use of the nurse's resources was significantly greater in the telemonitoring group (mean time at the reception was 48.7 minutes longer, and the number of nurse visits was 1.73 times greater, *P*<.001 and *P*<.001 respectively). There were more telephone contacts between the nurse and the telemonitoring patients (IRR=5.6 for nurse initiated contacts and IRR=1.63 for patient initiated contacts, *P*<.001 and *P*=.049 respectively). Typically the patients called the nurse for information to interpret the monitoring results (eg, the safe range for blood pressure) or because they wanted to make changes to their diuretic medication. The most frequent reason for nurse-induced calls was patients' non-adherence to self-monitoring: the nurse called patients to remind them to carry out and report the measurements. The number of unplanned visits to the Cardiology Outpatient Clinic was significantly bigger in the telemonitoring group (IRR=3.31, *P*<.001). The control group had on average one unplanned visit to the clinic while telemonitoring patients had 3-4 unplanned visits. The most common reason for unplanned visits was patients' concern about their worsening condition and the need to discuss it with the nurse. In some cases, patients visited physician reception if they needed immediate help. The reasons for phone calls and unplanned visits were based on nurse's notes. There was no difference in the use of physician resources: the number of visits and the time used at reception were similar between the study groups.

Depending on the patient's skills, 10-20 minutes was spent at the baseline visit for educating the patient to use the mobile phone app and blood pressure meter. During the follow-up period, only six telephone calls concerning purely technical problems took place. In three cases, the nurse called the patient during the first days to help them get started with the mobile app. The other three calls were initiated by the patient and were

XSL•FO
RenderX

caused due to failed Internet connections. All other contacts took place for medical reasons.

Table 4. The use of health care resources per patient during the study.

	Control group (n=47)	Telemonitoring group (n=46)	Effect size (95% CI)	*P* value[a]
Nurse time (minutes), mean (SD)	87 (35)	136 (43)	Mean difference 48.7 (32.5-64.8)	<.001
Number of visits to the nurse, mean (SD)	2.7 (1.0)	4.5 (2.2)	IRR=1.73 (1.38-2.15)	<.001
Physician time (minutes), mean (SD)	69 (23)	76 (34)	Mean difference 6.7 (6.0-18.6)	.340
Number of visits to the physician, mean (SD)	2.0 (0.8)	1.9 (0.9)	IRR=0.95 (0.71-1.28)	.738
Number of telephone contacts initiated by nurse, mean (SD)/[% patients with zero-count]	0.6 (0.9)/(57.4%)	3.0 (2.4)/(15.2%)	IRR=5.6 (3.41-7.63)	<.001
Number of telephone contacts initiated by patient, mean (SD)/[% patients with zero-count]	0.6 (1.5)/(72.3%)	2.3 (2.1)/(30.4%)	IRR=1.63 (0.999-2.66)	<.049
Number of unplanned visits to the Cardiology Outpatient Clinic, mean (SD)/[% patients with zero-count]	1.0 (1.5)/(46.8%)	3.7 (2.6)/(13%)	IRR=3.31 (2.15-5.09)	<.001

[a]Difference between groups.

Professional Experience

The HF nurse who was involved experienced telemonitoring as a valuable support to the current practice. She reported that the patients of the telemonitoring group took self-measurements more regularly and had internalized the importance of regular self-monitoring. Reception visits were more efficient, since no time was wasted on irrelevant issues. The nurse found that patients had taken their drugs more precisely, although no numerical evidence was collected. The nurse reported that both study groups were more curious about the ongoing study and that patients contacted her more frequently than prior to the study. The benefit that the nurse prioritized was the up-to-date data she received from the patients. The data also provided important support for physicians in their decisions about the patient's treatment, for example, in terms of adjustments to medication. A potential disadvantage that the nurse brought up was that the measurement data were input by the users; there was a possibility that some users sometimes sent false data by mistake or even intentionally. During the study, however, there were no signs of such problems. Automatic data transfer from monitoring devices would reduce the risk of erroneous data. The nurse responsible for the patients did not see any obstacles in adding telemonitoring as a part of their multidisciplinary care model.

Patient Experience

Of 46 patients, 44 (96%) responded to the user experience questionnaire. Almost all patients (95%, 42/44) found that making and reporting measurements with the mobile phone app was "very useful" or "quite useful". The automatic feedback they received after sending the measurements was found to be useful; in fact, 91% (40/44) of patients felt it was "very useful" or "quite useful". However, 9% (4/44) patients responded that they did not derive any benefit from the feedback. Two thirds (66%, 29/44) responded that the feedback helped them pay more attention to issues essential in the treatment of their disease. In fact, 91% (40/44) of patients responded that the feedback motivated them to take measurements and report them regularly. Just over a quarter of patients (27%, 12/44) reported that the feedback also gave them motivation to change their lifestyle.

Most of the patients accepted the home telemonitoring as part of their care routine. The adherence, calculated as a proportion of weekly submitted self-measurements, was 86% in weight reporting and 89% in blood pressure, heart rate, and symptom reporting. The median number of weight reports was 28 (interquartile range 23-33). The median number of blood pressure and symptom reports was 32 (interquartile range 27-43).

Post-Hoc Power Calculations

The post-hoc power was calculated using the Poisson model framework. Using the following definitions: exp(beta)=.8, base rate=0.03, total sample=97, mean exposure time=200 days, alpha=.05, R^2=0 (since the study group was the only predictor), distribution of group allocation= binomial with pi=0.5, the post-hoc power of 0.81 was obtained.

Discussion

Principal Results

This study evaluated whether a multidisciplinary care model would benefit from telemonitoring as an additional element in the care of heart failure patients, primarily in terms of reducing HF-related hospital days. We found that the telemonitoring-assisted care approach led to increased use of health care resources while showing no quantified improvement in the patients' condition. There was no difference in the number of HF-related hospital days, which was the primary outcome.

XSL•FO
RenderX

However, patients and health care providers reacted positively to telemonitoring. Patients' adherence to the weekly reporting plan was close to 90%, which is high in a population with a severe chronic condition.

The increased use of health care resources was primarily seen in the nurse's workload. The telemonitoring group had an increased number of visits to the nurse reception, a longer time spent at the reception, and more frequent telephone contacts with the nurse. Similarly, in the studies by Cleland et al [16] and Wade et al [19], home telemonitoring was associated with frequent patient contacts including home and office visits, telephone contacts, and emergency visits. However, in our study there was neither increased need for physician consultation nor increased number of visits to the emergency department. Despite the increased workload, the nurse found the increased number of contacts with patients to be a positive change. In her experience, telemonitoring invoked the patients' interest in their condition and raised questions that resulted in contacts. The nurse also felt that the control group patients were more active after their enrollment in the study. Patients' increased curiosity was not reflected in a lower number of HF-related hospitalizations, but we can speculate whether low death rates were associated with this. During the 6-month follow-up, we found no deaths in either of the study groups, which is unexpected since the mortality rate at 30 days after hospital admission is 11% [27], rising to 30% during the first year [3].

When implementing telemonitoring in the care process, the increased workload of care professionals needs to be accounted for. Patients' increased awareness of their disease is likely to increase contacts. Patients need help in interpreting the monitoring results, and they seek individual advice in order to manage their disease and maintain their enthusiasm. This kind of activity may lead to positive health outcomes during a longer follow-up period than was the case in the present study. It should be carefully considered whether the current resources are able to handle the increased demand, or whether additional personnel should be hired. As Chaundry et al [14] concluded, a telemonitoring strategy would be more effective if embedded in cardiology practices with a greater organizational capacity to implement it. To lessen the increased workload of health care professionals, the potential of active assistance technology is worthy of consideration. Such technologies include highly sophisticated automatic messaging systems providing personalized guidance to patients with minimum involvement of health care personnel. Promising results with active assistance technology have been reported in the care of diabetes patients [28].

In this study, telemonitoring was linked to more individualized care in terms of the pharmacological therapy of HF patients. This shows an important aspect, since optimal pharmacological management reduces morbidity and mortality, but it is complex and objective guidelines are lacking [29]. Significantly more changes in medication regimen were made in the telemonitoring group—medication was increased for 17% of these patients whereas in the control group the corresponding percentage was 4%. In addition, all reductions in medication were done for telemonitoring patients. Reductions were applied to Furosemide, which is a diuretic, indicating successful management of fluid retention. Whether medication changes were the result of self-measurement data that telemonitoring patients provided or through their increased self-care or both cannot be confirmed with these data. The frequency of which the measurements were done may alone not be sufficient to constitute the medication changes, but at least the intervention opened patients' eyes and raised discussion concerning their medication.

Our negative finding regarding the hospitalization rate is in line with findings in studies [14-20]. However, several studies do show evidence in favor of telemonitoring in the care of HF patients. This brings up the challenge in providing telehealth for the right patients in the right context. In the TIM-HF study, a subgroup analysis revealed that patients with lower depression scores had significantly lower hospitalization and mortality [30]. Comorbidities such as chronic obstructive pulmonary disease, chronic kidney disease, and anemia may have a negative effect by blurring the signal from the monitored variables and thus lowering their predictive value [31]. Furthermore, in their review of telemonitoring for chronic diseases, Pare et al [32] concluded that the beneficial effects of telemonitoring are more consistent in pulmonary and cardiac studies than in diabetes and hypertension. In the tele-HF study [14], the authors concluded that none of the participant characteristics including age, sex, race, LVEF, and NYHA class identified a group in which telemonitoring was more effective. A similar conclusion was drawn in [30] in terms of LVEF, gender, age, or NYHA.

We outline four factors that may be associated with unchanged HF-related hospitalizations rates. First, as multidisciplinary care was part of the care practice at the Cardiology Outpatient Clinic of HUCH, all the patients including the control group received high standard interactive care, and some were already used to home self-monitoring. Second, the study was carried out in the Helsinki area where patients live a short distance from health services. Patients were able to visit the clinic easily without great effort. We found that the most common reason for unplanned visits was that the patient wanted to discuss with the nurse face-to-face the signs of deterioration and worsening condition. Home telemonitoring may be more beneficial when applied in rural areas where patients do not have direct access to health care. Third, our study population was relatively young, and medication for all patients was optimized during the baseline visit. Ejection fraction was on average 28%, which is higher than in the TEN-HMS study [16] of high-risk HF-patients. Telemonitoring may be more efficient among patients with poor prognosis. In the TEN-HMS study, which found home telemonitoring associated with improved survival, patients were older, had severe cardiac dysfunction, were recently hospitalized, and had high mortality rates. Finally, the follow-up time was possibly too short. Improved self-care may be realized as a lower number of hospitalizations after a time interval longer than 6 months.

Limitations

Post-hoc calculations were conducted based on the Poisson model framework resulting in a power of 0.81, which was less than was determined in initial power calculations. Considering the fact that the 95% confidence interval for the IRR ranged from 0.525 to 1.256, we do not expect that there was a true

XSL•FO
RenderX

difference in the number of HF-related hospital days between the study groups, although we did not reach the level of 0.1 for the type II error. However, the predicted difference of 3 days was overestimated since the number of hospital days was 1.4 in the control group. In addition, we note that we conducted multiple hypothesis testing, which increases the probability of falsely rejecting the null hypothesis. However, the statistically significant findings that were seen in the use of health care resources were consistent in several variables supporting each other.

The usage of the nurse's time was somewhat biased. The time consumed at the baseline visit for the delivery of telemonitoring technology to the patients was counted as time spent by the nurse. Also, when technical problems emerged, patients contacted the nurse. The time used at baseline visit was 10-20 minutes per patient. During the monitoring period, only six contacts were made with the nurse due to technical problems. Therefore, it can be concluded that the time spent on technical issues was marginal and that the increased use of nurse's resources by telemonitoring patients took place due to medical reasons. Technical issues did not increase the required time to an extent that would lead to significant overestimation. An additional source of bias is the fact that monitoring took place under control of only one research nurse and the professional experience was based only on her interview. Consequently, it is not possible to draw general conclusions on the attitudes of health care professionals on monitoring.

Conclusions

In the Heart at Home study, we found that home telemonitoring was not efficient to support the multidisciplinary care approach in terms of reducing the number of HF-related hospital days or outpatient visits or improving patients' clinical condition. The telemonitoring increased significantly the nurse's workload by increasing the number of reception visits and the number of telephone contacts. The increased workload should be carefully considered when implementing telemonitoring in the care of HF patients. Extra work is required on top of the multidisciplinary care approach. To lessen the increased workload of health care professionals, the potential of active assistance technology is worthy of further consideration to respond to patients' queries and to keep them motivated.

Acknowledgments

We wish to thank the personnel of the Cardiology Outpatient Clinic of HUCH, and in particular the research nurse Leila Partanen, for carrying out the clinical part of the study. The Finnish Funding Agency for Technology and Innovation and VTT Technical Research Centre of Finland are acknowledged for funding the study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT EHEALTH checklist V1.6.2 [25].

[PDF File (Adobe PDF File), 999KB - jmir_v16i12e282_app1.pdf]

References

1. McMurray JJV, Stewart S. The burden of heart failure. Eur Heart J Suppl 2002;4(suppleD):307-333.
2. Lang CC, Mancini DM. Non-cardiac comorbidities in chronic heart failure. Heart 2007 Jun;93(6):665-671 [FREE Full text] [doi: 10.1136/hrt.2005.068296] [Medline: 16488925]
3. Chen J, Normand SL, Wang Y, Krumholz HM. National and regional trends in heart failure hospitalization and mortality rates for Medicare beneficiaries, 1998-2008. JAMA 2011 Oct 19;306(15):1669-1678 [FREE Full text] [doi: 10.1001/jama.2011.1474] [Medline: 22009099]
4. Krumholz HM, Merrill AR, Schone EM, Schreiner GC, Chen J, Bradley EH, et al. Patterns of hospital performance in acute myocardial infarction and heart failure 30-day mortality and readmission. Circ Cardiovasc Qual Outcomes 2009 Sep;2(5):407-413 [FREE Full text] [doi: 10.1161/CIRCOUTCOMES.109.883256] [Medline: 20031870]
5. Desai AS, Stevenson LW. Rehospitalization for heart failure: predict or prevent? Circulation 2012 Jul 24;126(4):501-506 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.112.125435] [Medline: 22825412]
6. Stewart S. Financial aspects of heart failure programs of care. Eur J Heart Fail 2005 Mar 16;7(3):423-428. [doi: 10.1016/j.ejheart.2005.01.001] [Medline: 15718184]
7. Stewart S, MacIntyre K, Hole DJ, Capewell S, McMurray JJ. More 'malignant' than cancer? Five-year survival following a first admission for heart failure. Eur J Heart Fail 2001 Jun;3(3):315-322. [Medline: 11378002]
8. McAlister FA, Stewart S, Ferrua S, McMurray JJ. Multidisciplinary strategies for the management of heart failure patients at high risk for admission: a systematic review of randomized trials. J Am Coll Cardiol 2004 Aug 18;44(4):810-819 [FREE Full text] [doi: 10.1016/j.jacc.2004.05.055] [Medline: 15312864]
9. van der Wal MH, Jaarsma T, van Veldhuisen DJ. Non-compliance in patients with heart failure; how can we manage it? Eur J Heart Fail 2005 Jan;7(1):5-17. [doi: 10.1016/j.ejheart.2004.04.007] [Medline: 15642526]

XSL•FO
RenderX

10. Michalsen A, König G, Thimme W. Preventable causative factors leading to hospital admission with decompensated heart failure. Heart 1998 Nov;80(5):437-441 [FREE Full text] [Medline: 9930040]
11. Klersy C, De Silvestri A, Gabutti G, Regoli F, Auricchio A. A meta-analysis of remote monitoring of heart failure patients. J Am Coll Cardiol 2009 Oct 27;54(18):1683-1694 [FREE Full text] [doi: 10.1016/j.jacc.2009.08.017] [Medline: 19850208]
12. Inglis SC, Clark RA, McAlister FA, Stewart S, Cleland JG. Which components of heart failure programmes are effective? A systematic review and meta-analysis of the outcomes of structured telephone support or telemonitoring as the primary component of chronic heart failure management in 8323 patients: Abridged Cochrane Review. Eur J Heart Fail 2011 Sep;13(9):1028-1040. [doi: 10.1093/eurjhf/hfr039] [Medline: 21733889]
13. Inglis SC, Clark RA, McAlister FA, Ball J, Lewinter C, Cullington D, et al. Structured telephone support or telemonitoring programmes for patients with chronic heart failure. Cochrane Database Syst Rev 2010(8):CD007228. [doi: 10.1002/14651858.CD007228.pub2] [Medline: 20687083]
14. Chaudhry SI, Mattera JA, Curtis JP, Spertus JA, Herrin J, Lin Z, et al. Telemonitoring in patients with heart failure. N Engl J Med 2010 Dec 9;363(24):2301-2309 [FREE Full text] [doi: 10.1056/NEJMoa1010029] [Medline: 21080835]
15. Koehler F, Winkler S, Schieber M, Sechtem U, Stangl K, Böhm M, Telemedical Interventional Monitoring in Heart Failure Investigators. Impact of remote telemedical management on mortality and hospitalizations in ambulatory patients with chronic heart failure: the telemedical interventional monitoring in heart failure study. Circulation 2011 May 3;123(17):1873-1880 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.111.018473] [Medline: 21444883]
16. Cleland JG, Louis AA, Rigby AS, Janssens U, Balk AH, TEN-HMS Investigators. Noninvasive home telemonitoring for patients with heart failure at high risk of recurrent admission and death: the Trans-European Network-Home-Care Management System (TEN-HMS) study. J Am Coll Cardiol 2005 May 17;45(10):1654-1664 [FREE Full text] [doi: 10.1016/j.jacc.2005.01.050] [Medline: 15893183]
17. Mortara A, Pinna GD, Johnson P, Maestri R, Capomolla S, La Rovere MT, et al. Home telemonitoring in heart failure patients: the HHH study (Home or Hospital in Heart Failure). Eur J Heart Fail 2009 Mar;11(3):312-318 [FREE Full text] [doi: 10.1093/eurjhf/hfp022] [Medline: 19228800]
18. Takahashi PY, Pecina JL, Upatising B, Chaudhry R, Shah ND, Van Houten H, et al. A randomized controlled trial of telemonitoring in older adults with multiple health issues to prevent hospitalizations and emergency department visits. Arch Intern Med 2012 May 28;172(10):773-779 [FREE Full text] [doi: 10.1001/archinternmed.2012.256] [Medline: 22507696]
19. Wade MJ, Desai AS, Spettell CM, Snyder AD, McGowan-Stackewicz V, Kummer PJ, et al. Telemonitoring with case management for seniors with heart failure. Am J Manag Care 2011 Mar;17(3):e71-e79 [FREE Full text] [Medline: 21504262]
20. Seto E, Leonard KJ, Cafazzo JA, Barnsley J, Masino C, Ross HJ. Mobile phone-based telemonitoring for heart failure management: a randomized controlled trial. J Med Internet Res 2012;14(1):e31 [FREE Full text] [doi: 10.2196/jmir.1909] [Medline: 22356799]
21. Steventon A, Bardsley M, Billings J, Dixon J, Doll H, Hirani S, Whole System Demonstrator Evaluation Team. Effect of telehealth on use of secondary care and mortality: findings from the Whole System Demonstrator cluster randomised trial. BMJ 2012;344:e3874 [FREE Full text] [Medline: 22723612]
22. Cartwright M, Hirani SP, Rixon L, Beynon M, Doll H, Bower P, Whole Systems Demonstrator Evaluation Team. Effect of telehealth on quality of life and psychological outcomes over 12 months (Whole Systems Demonstrator telehealth questionnaire study): nested study of patient reported outcomes in a pragmatic, cluster randomised controlled trial. BMJ 2013;346:f653 [FREE Full text] [Medline: 23444424]
23. Henderson C, Knapp M, Fernández J, Beecham J, Hirani SP, Cartwright M, Whole System Demonstrator evaluation team. Cost effectiveness of telehealth for patients with long term conditions (Whole Systems Demonstrator telehealth questionnaire study): nested economic evaluation in a pragmatic, cluster randomised controlled trial. BMJ 2013;346:f1035 [FREE Full text] [Medline: 23520339]
24. Steventon A, Bardsley M, Billings J, Dixon J, Doll H, Beynon M, et al. Effect of telecare on use of health and social care services: findings from the Whole Systems Demonstrator cluster randomised trial. Age Ageing 2013 Jul;42(4):501-508 [FREE Full text] [doi: 10.1093/ageing/aft008] [Medline: 23443509]
25. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
26. Vuong Q. Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica 1989;57:307-333. [doi: 10.2307/191255]
27. Bueno H, Ross JS, Wang Y, Chen J, Vidán MT, Normand SL, et al. Trends in length of stay and short-term outcomes among Medicare patients hospitalized for heart failure, 1993-2006. JAMA 2010 Jun 2;303(21):2141-2147 [FREE Full text] [doi: 10.1001/jama.2010.748] [Medline: 20516414]
28. Orsama AL, Lähteenmäki J, Harno K, Kulju M, Wintergerst E, Schachner H, et al. Active assistance technology reduces glycosylated hemoglobin and weight in individuals with type 2 diabetes: results of a theory-based randomized trial. Diabetes Technol Ther 2013 Aug;15(8):662-669. [doi: 10.1089/dia.2013.0056] [Medline: 23844570]

XSL•FO
RenderX

29. Troughton RW, Frampton CM, Yandle TG, Espiner EA, Nicholls MG, Richards AM. Treatment of heart failure guided by plasma aminoterminal brain natriuretic peptide (N-BNP) concentrations. Lancet 2000 Apr 1;355(9210):1126-1130. [Medline: 10791374]
30. Koehler F, Winkler S, Schieber M, Sechtem U, Stangl K, Böhm M, et al. Telemedicine in heart failure: pre-specified and exploratory subgroup analyses from the TIM-HF trial. Int J Cardiol 2012 Nov 29;161(3):143-150. [doi: 10.1016/j.ijcard.2011.09.007] [Medline: 21982700]
31. Anker SD, Koehler F, Abraham WT. Telemedicine and remote management of patients with heart failure. Lancet 2011 Aug 20;378(9792):731-739. [doi: 10.1016/S0140-6736(11)61229-4] [Medline: 21856487]
32. Paré G, Jaana M, Sicotte C. Systematic review of home telemonitoring for chronic diseases: the evidence base. J Am Med Inform Assoc 2007;14(3):269-277 [FREE Full text] [doi: 10.1197/jamia.M2270] [Medline: 17329725]

Abbreviations

ACE.-I: angiotensin converting enzyme inhibitor
EHFSBS: European Heart Failure Self-Care Behaviour Scale
HF: heart failure
HUCH: Helsinki University Central Hospital
IRR: incidence rate ratio
LVEF: left ventricular ejection fraction
NT-proBNP: N-terminal of the prohormone brain natriuretic peptide
NYHA: New York Heart Association
WSD: Whole Systems Demonstrator

Edited by G Eysenbach; submitted 30.06.14; peer-reviewed by J Cleland, P Kastner, L Ning; comments to author 31.07.14; revised version received 10.09.14; accepted 05.11.14; published 11.12.14.

Please cite as:
Vuorinen AL, Leppänen J, Kaijanranta H, Kulju M, Heliö T, van Gils M, Lähteenmäki J
Use of Home Telemonitoring to Support Multidisciplinary Care of Heart Failure Patients in Finland: Randomized Controlled Trial
J Med Internet Res 2014;16(12):e282
URL: http://www.jmir.org/2014/12/e282/
doi:10.2196/jmir.3651
PMID:25498992

©Anna-Leena Vuorinen, Juha Leppänen, Hannu Kaijanranta, Minna Kulju, Tiina Heliö, Mark van Gils, Jaakko Lähteenmäki. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 11.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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

Tutorial

Web-Based Assessments of Physical Activity in Youth: Considerations for Design and Scale Calibration

Pedro F Saint-Maurice[1,2], PhD; Gregory J Welk[1], PhD

[1]Iowa State University, Department of Kinesiology, Ames, IA, United States

[2]University of Minho, School of Psychology, CIPsi, Braga, Portugal

Corresponding Author:
Pedro F Saint-Maurice, PhD
Iowa State University
Department of Kinesiology
283 Forker Building
Ames, IA, 50010
United States
Phone: 1 515294 8009
Fax: 1 515294 8740
Email: pedrosm@iastate.edu

Abstract

This paper describes the design and methods involved in calibrating a Web-based self-report instrument to estimate physical activity behavior. The limitations of self-report measures are well known, but calibration methods enable the reported information to be equated to estimates obtained from objective data. This paper summarizes design considerations for effective development and calibration of physical activity self-report measures. Each of the design considerations is put into context and followed by a practical application based on our ongoing calibration research with a promising online self-report tool called the Youth Activity Profile (YAP). We first describe the overall concept of calibration and how this influences the selection of appropriate self-report tools for this population. We point out the advantages and disadvantages of different monitoring devices since the choice of the criterion measure and the strategies used to minimize error in the measure can dramatically improve the quality of the data. We summarize strategies to ensure quality control in data collection and discuss analytical considerations involved in group- vs individual-level inference. For cross-validation procedures, we describe the advantages of equivalence testing procedures that directly test and quantify agreement. Lastly, we introduce the unique challenges encountered when transitioning from paper to a Web-based tool. The Web offers considerable potential for broad adoption but an iterative calibration approach focused on continued refinement is needed to ensure that estimates are generalizable across individuals, regions, seasons and countries.

(J Med Internet Res 2014;16(12):e269) doi:10.2196/jmir.3626

KEYWORDS

measurement; questionnaire; Youth Activity Profile

Introduction

Physical activity behaviors can be assessed with a variety of techniques, but each has inherent limitations [1]. Researchers have increasingly used objective monitoring devices to capture ambulatory movement [2], but there is a fundamental need to also improve the utility of self-report instruments that can be more effectively deployed in Web-based applications in a cost effective way. The sophistication (and accuracy) of objective monitoring devices has increased dramatically in recent years through the use of new technologies [3,4] and incremental advances in calibration methodologies [5-11]. Surprisingly, relatively few efforts have been made to improve the utility of more subjective, self-report measures. The *Journal of Physical Activity & Health* published the outcomes of the Measurement of Active and Sedentary Behaviors: Closing the Gaps in Self-Report Methods meeting sponsored by the National Institutes of Health that addressed a series of best practices and challenges encountered when working with self-report tools for physical activity [12]. One of the limitations of self-report instruments is a general lack of accuracy; however, the use of robust calibration procedures and implementation of Web-based applications offer considerable promise for improving the validity and utility of self-report measures [12-20]. A specific advantage is that algorithms from these methods can be directly embedded within existing Web-based applications to streamline collection and reporting of data. For optimal effectiveness, these

XSL•FO
RenderX

models must be developed, tested, and then refined through an iterative process that enables enhancements to be directly incorporated into the assessment.

Measurement error models have been used to understand error in self-reported data on food intake [21-23] and physical activity behavior [22,24,25], but this work is still in early phases. An advantage in physical activity research is that there are available criterion measures from objective monitoring devices that can be used to scale the information provided on the self-report forms. However, there are a number of challenges involved in effectively linking self-report data to objective data for calibration. There are a number of other design considerations that also influence the ability to calibrate self-report data. The present paper describes key principles and design considerations needed for effective calibration of self-reported physical activity data. This paper also addresses the potential application of online self-report tools for large-scale assessment. Examples are based on work we have done to calibrate the Youth Activity Profile (YAP), a promising online self-report measure of physical activity in youth. Although some issues are specific to this instrument, the design principles, methods of calibration, and Web applications would apply to other surveys and populations.

The YAP provides a good example for illustrating calibration design and online applications because it was developed specifically with these goals in mind. It was designed for use in school settings to make it possible to obtain accurate data from youth while also providing a valuable educational experience. The YAP uses a segmented day approach and context-related events to facilitate recall by youth. It captures the key dimensions of physical activity needed for calibration (eg, frequency, type, volume), and includes separate components to capture the context of physical activity (eg, in-school vs. out-of-school) to facilitate education and promotion.

We have established the Youth Physical Activity Measurement Study (YPAMS) to facilitate continued development and refinement of the YAP. This paper first describes the key components of the YAP because examples are based on our experience with this model. We then summarize the design features and calibration principles that are important for this type of calibration work. We conclude with a discussion of the unique considerations required for the calibration of an online version of calibrated tools in the final section.

Development of the Youth Activity Profile

The YAP is a self-report instrument designed to capture physical activity and sedentary behavior in youth. It was based conceptually on the time-based structure of the established Physical Activity Questionnaire for Children (PAQ-C) and a related tool for adolescents (PAQ-A) [26-29]. We identified the Physical Activity Questionnaire (PAQ) as a promising platform because it was consistent with established recommendations for recall instruments (ie, uses short-term recall periods, uses contextual questions, stimulates episodic memories, and asks about overall moderate-to-vigorous physical activity, MVPA). In addition, the PAQ is a short form that can provide direct comparisons among youth of different age groups. The PAQ takes less than 15-20 minutes to be completed; therefore, this instrument is very practical. All these might explain the overall good psychometric properties of this instrument [26-28,30-33].

The YAP was designed to be a self-administered 7-day (previous week) recall questionnaire suitable for use with children in grades 4 to 12. The structure for some items maintains the conceptual idea of the PAQ, but the individual items were changed to improve calibration. Additional items were added and a whole category of sedentary time was developed. The YAP includes a total of 15 items divided into 3 sections: (1) activity at school, (2) activity outside of school, and (3) sedentary behaviors. The final version of the YAP was pilot tested and cognitive interviews (in laboratory) were performed with similarly aged participants to refine the final items. The final version of the YAP is available in Multimedia Appendix 1.

The scoring procedures used in the YAP are similar to those used in the original PAQ; however, each YAP section was developed to be scored independently (ie, items for each dimension are averaged to reflect the composite score of the respective dimension). A higher composite score for a given dimension score would reflect a higher expected activity levels/sedentary time at that same dimension. Preliminary calibration of the YAP has been conducted using a paper version (phase I and II). However, an online version of the YAP has been developed and will be used for future refinement (phase III) (Figure 1). Online versions of existent physical activity questionnaires can provide considerable advantages for both education (school) and research applications because data can be entered more easily and processed automatically to enable feedback.

XSL•FO
RenderX

Figure 1. Screen capture of the online (top) and game (bottom) version of the Youth Activity Profile.



Design Considerations

Overview

In this paper, *calibration* refers to the process of computing adjusted minutes of objectively measured physical activity from a raw self-reported physical activity index [34]. A number of factors must be considered when setting up a calibration study and each decision has ripple effects on other aspects of the design. We first describe the overall calibration design and follow with the importance of selecting an appropriate criterion measure and then describe key considerations, such as participant selection, sample size determination, measurement error, and validity of calibration algorithms for online applications. At each of these sections, we provide some contextual information of the concept/issue being discussed and then provide a practical example based on our experience with the Youth Physical Activity Measurement Study (YPAMS).

Calibration Design

Background

To enhance calibration, the structure and items from self-report tools should be designed to link objective data to the subjective responses. Each question should capture a discrete period of time in which there are specific opportunities to be physically active. Descriptive studies provide documentation of settings or periods of time that capture significant amounts of physical activity [35,36]. If similar periods are identified and included in the self-report tool, it is possible to link each period to corresponding data from the activity monitor and pursue individual calibration. See sample temporal linkage provided in Table 1.

YPAMS Example

There are 5 items in the YAP that capture periods at school (ie, walk/bike to school, at recess, during physical education (PE) class, at lunch, walk/bike from school) and 5 items that capture periods at home (before school, after school, evening, Saturday and Sunday). The individual time segments are directly matched with corresponding time periods from the accelerometer (see Figure 2 for examples of student activity during PE, recess, and Sunday). The calibration process averages responses over the available number of days (and then across participants) to obtain stable estimates for each of the segments. Some time segments are standardized across days (eg, Before school, After School etc…); however, other segments such as PE, recess and lunch can vary across days so this flexibility was built into the coding.

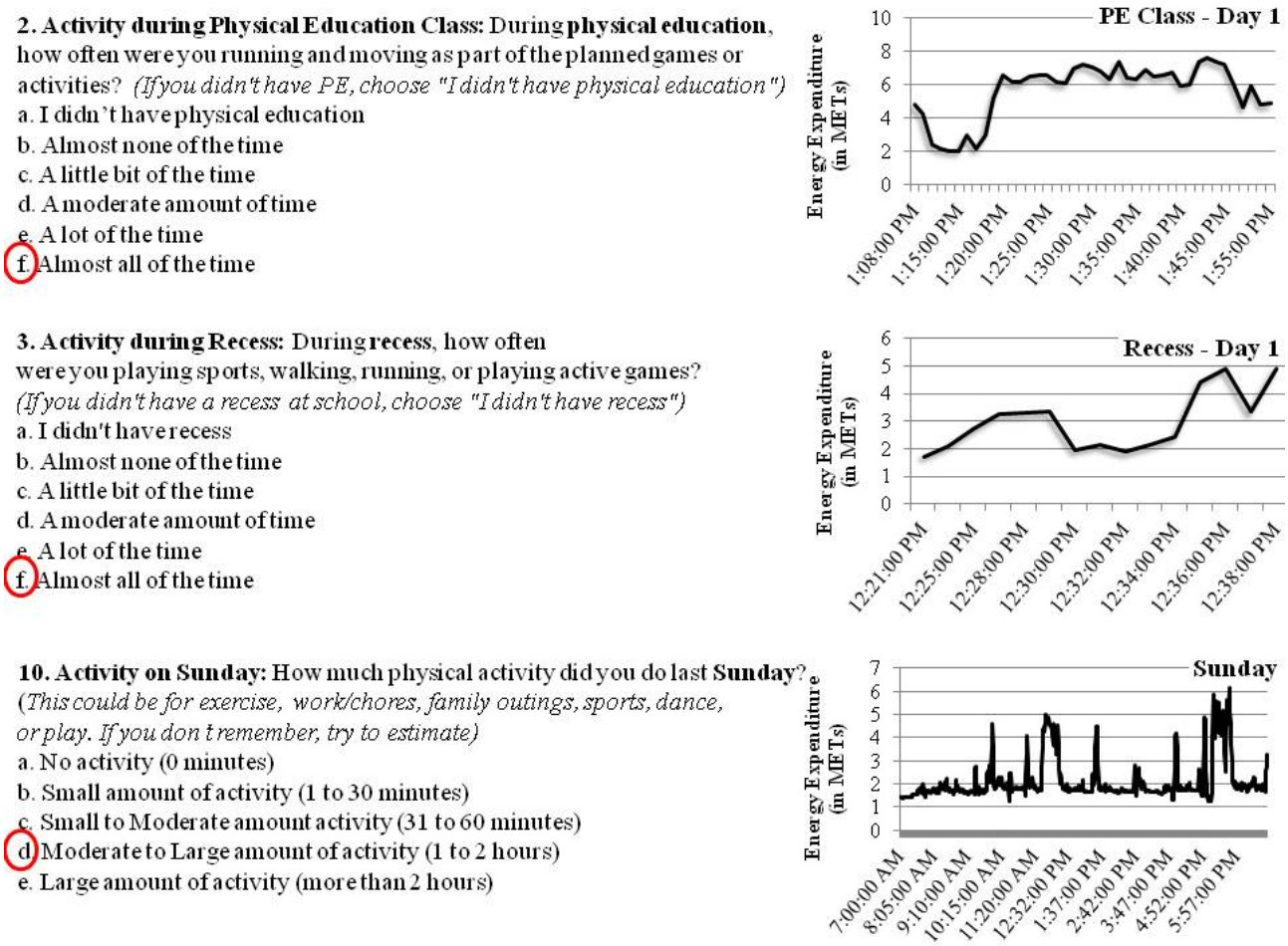
XSL•FO
RenderX

Table 1. Weekly schedule used to process segmented accelerometer data.

Window	Date	Individualized time	Start time[a]	End time[a]
Before school	Every day	Yes	60 min before start time for trans to school	Start time for trans to school
Transportation to school	Every day	Yes	30 min before start time for school	Start time for school
Recess	Provided	Yes		
Physical Education	Provided	Yes		
Lunch	Every day	Yes		
Transportation from school	Every day	Yes	End time for school	30 min after end time for school
After school	Every day	Yes	End time for trans from school	6:00 pm
Evening	Every day	No	6:00 pm	10:00 pm
Saturday	Saturday	No	7:00 am	10:00 pm
Sunday	Sunday	No	7:00 am	10:00 pm

[a] "Start" and "end" school time was obtained from schools (eg, 8:15 am-3:30 pm).

Figure 2. Examples of energy expenditure values measured by accelerometer for discrete time segments captured by the YAP. The data presented are from a middle school female participant enrolled in the YPAMS.



XSL•FO
RenderX

Selection of the Criterion Measure

Background

There are many options available for capturing objective information about physical activity, but each has inherent advantages and disadvantages. The most widely used accelerometry-based activity monitor is the Actigraph, but a limitation of this device is that it is not possible to directly detect if a participant is wearing the device. Numerous studies have been conducted with the Actigraph, but there is still a need to develop better algorithms to capture EE [37,38] or to detect wear time [39,40].

YPAMS Example

For our YAP calibration work, we selected the SenseWear Armband (BodyMedia, Inc, Pittsburgh, PA, USA) as our criterion measure of physical activity. We conducted numerous studies with various accelerometry-based monitors and found the SenseWear Armband offered a number of advantages for use in calibration studies. The SenseWear Armband is a wireless pattern-recognition device that integrates motion sensor data with a variety of heat-related sensors and demographic variables to estimate the EE [41]. The multisensor nature of the monitor provides advantages over traditional accelerometry-based monitors. The heat-related sensors, for example, provide a better indicator of work (and EE) for nonlocomotor tasks and activities of daily living [42]. An additional advantage of the SenseWear Armband for field-based research is that it automatically detects periods of time in which the accelerometer is not worn (ie, removed by the participants). This is still a common source of error in physical activity studies that use more standard measures of physical activity [39].

Characteristics of the Study Population

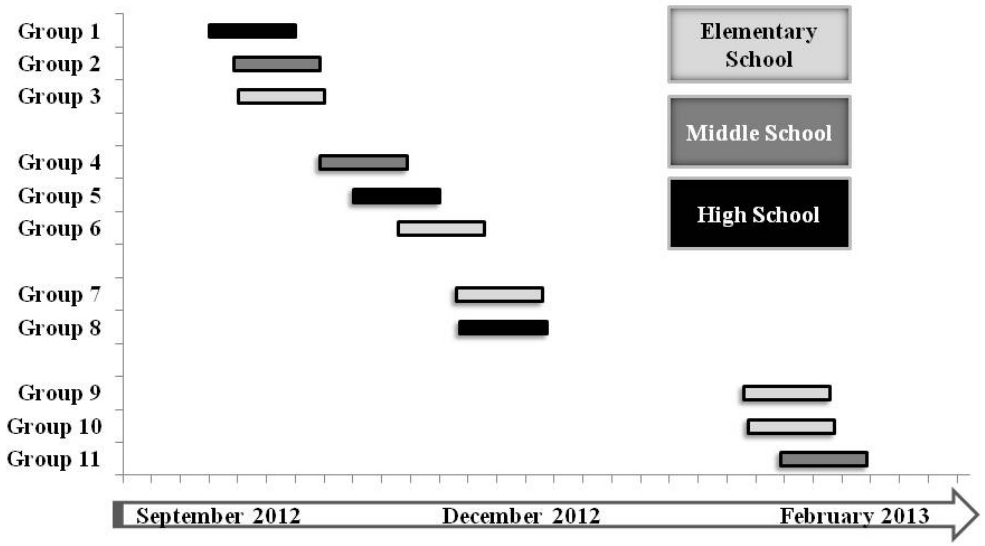
Background

The nature of the study population should be taken into account when developing recruitment and retention strategies. It is important to ensure that the sample population is representative of the larger population and that there is sufficient variability within the sample. Variability in levels of physical activity is especially important in calibration design to ensure that it can predict both low and high levels of physical activity. The calibration of a scale is also enhanced if the observed exposure values (MVPA from the SenseWear Armband) are normally distributed within each level of the covariates in the model being considered (eg, age and gender). Although controlling for covariates is necessary, one would also need to maximize the variability in your exposure variable. This task can be challenging because calibration of self-report tools are designed to capture real life physical activity patterns as opposed to common calibration activity monitors studies that are conducted in controlled environments and, therefore, allow activities to be manipulated [43,44]. One possible strategy to artificially manipulate individual activity patterns while controlling for age and gender is to recruit individuals with different activity levels.

YPAMS Example

In our calibration work, we enhanced variability in activity levels by recruiting a diverse sample and by collecting activity data across seasons (eg, winter and summer). Season has been shown to influence physical activity patterns in youth [45], so the advantage of this approach is that the calibrations would capture this inherent variability and be more generalizable. The YPAMS design also counterbalanced the dates of data collection by age group (elementary, middle, and high school grades). For example, at each season, data are collected in elementary, middle, and high school participant groups composed by an approximately similar number of boys and girls, and in a counterbalanced order. Figure 3 provides an illustration of the data collection map for a full year of the YPAMS.

Figure 3. Gant chart illustrating the timeline and design for data collection of the preliminary work. Each rectangle represents a classroom grade full protocol timeline. Thick bars on the horizontal axis represent 1 calendar week. Each collection took approximately 3 weeks.



XSL•FO
RenderX

Sample Size and Recruitment Issues

Background

Sample size calculations for physical activity calibration studies should account for important predictors of physical activity in youth. Many correlates of physical activity have been identified, but age and gender are perhaps the most important factors to consider in calibration studies. These 2 factors are known to be related with objectively measured physical activity and, therefore, will affect statistical power of the questionnaire being calibrated. Sample size estimations for calibration can be computed using multiple regression with random factors and by including 3 predictors for the full model (age, gender, and questionnaire score) at a specific alpha and power. The expected variance accounted for with the full model would depend on the nature of the assessment, but this can be determined by conducting a systematic review of the literature. In addition to statistical power, researchers should also consider typical compliance rates with activity monitoring studies and account for noncompliance by overrecruiting participants.

YPAMS Example

In our preliminary work of the YPAMS, we used population level variances of 0.35 for a full model with the 3 predictors described previously and 0.25 for the reduced model. These variances were defined after reviewing the literature on the agreement between self-reports and accelerometry in youth. We had compliance rates of approximately 89% and, on average, participants (including compliant and noncompliant) wore the accelerometer during 71% of awake time (ie, between 7:00 am and 10:00 pm). We learned that younger children are more willing to participate in these studies and that they tend to be more compliant compared with their older peers. We have incorporated these lessons in our YPAMS study to promote adherence and compliance with our protocol. Direct collaboration with school districts and PE teachers has facilitated recruitment and retention in our school-based testing.

Measurement Error in the Criterion (Objective) Measure

Background

In the context of linear regression, measurement error can lead to a variety of flawed outcomes that can go from the attenuation of the relations between the exposure variable and health-related outcomes [46] to reversed effects when error-free measures are used [47]. The same concern applies to physical activity and, in particular, to calibration studies that have to rely on non-error-free gold-standard measures [48]. Measures obtained from activity monitor tools can lead to both random and systematic error and include (1) malfunction and handling of activity monitors, (2) day-to-day variability in activity patterns, and (3) periods of nonwear time. These sources of error can be attenuated with standardization of procedures and by scheduling periodic calibration checks of the monitors being used. Another key to reducing error is to collect data across multiple days and to define strict criteria to evaluate compliance in monitoring. Commonly used data reduction procedures include requiring valid activity counts for a minimum period of time (eg, 10 hours per day and/or 3 days per week). However, researchers can minimize the loss of data by having participants record descriptive information about nonwear bouts (eg, sports events). Data can be imputed using EE values (ie, METs) available for this purpose [49-51]. The standard MET values used for imputing missing periods were not developed to account for individual differences in the activities being considered, but the inclusion of these values avoids the alternative problem of deleting cases / time periods in which participants were not able to wear the monitors and can capture important activity that otherwise would be dismissed [52,53]. Overall, the choice of data reduction procedures will have an impact on the criterion measure outcome [54,55].

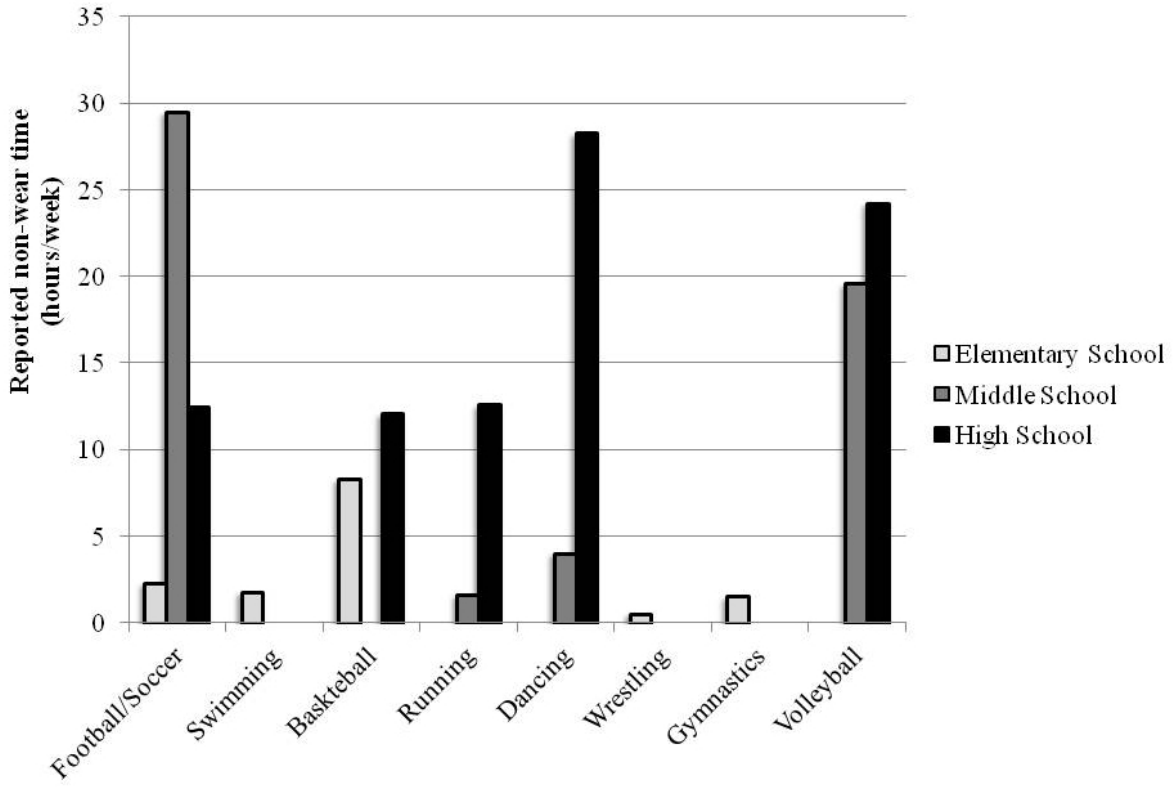
YPAMS Example

In our work, we have minimized noncompliance by sending information home to parents and by calling participants early in the week to remind them about the importance of wearing the monitor. Participants were also required to have valid EE values during at least 70% of the time and at least 3 valid periods for each time window of interest (exception was made for PE class, Saturday, and Sunday segments in which participants are only required to have at least 1 valid segment). Information about nonwear time was obtained from a daily activity log and these data were used to impute standardized MET values from the compendium of physical activities [49-51]. Figure 4 provides an illustration of nonreported activity in the YPAMS study.

By imputing missing time periods (due to documented nonwear times), we have been able to maximize the available sample. We noted that nonwear time was more prevalent in older students, possibly reflecting the changes in daily physical activity patterns as children get older.

XSL•FO
RenderX

Figure 4. Reported nonwear time across 7 days obtained from individual logs. Data presented are from a subset of participants enrolled in the YPAMS. Results are based on a sample of 87 elementary, 27 middle school, and 29 high school students that used a SenseWear Armband for 7 consecutive days. There were 2 participants without accelerometer data who provided records of nonwear time.



Measurement Error in the Calibrated (Subjective) Measure

Background

It is important to consider sources of error in the self-report measure so that they can be quantified and modeled appropriately. Self-report tools are known to be susceptible to different sources of error, but the majority of the error can be attributed to the nature of the instrument and the individual's subjectivity when interpreting and completing the questionnaire. For effective calibration, it is important to try to minimize error from these (and other) sources. It is important to ensure that the questions are appropriately worded and easy to understand. A formalized approach using cognitive interviews can help to identify issues with the survey before it is finalized. Error derived from the individual subjectivity can be minimized by adding recall probes and conducting quality checks after the questionnaire is completed. These 2 factors are particularly important in younger populations. Efforts to collect higher quality data from the survey will reduce error and improve calibration accuracy.

YPAMS Example

In our calibration work, we performed cognitive interviews by having individual face-to-face interviews with participants and asking them to think aloud as they were completing the YAP. Additionally, during our field data collection, a trained research member verified each child's questionnaire once the questionnaire was completed. Children were asked to provide more specific information (eg, exact location, exact day, or time of the day) about the behavior reported. In some cases, this process helped verify that participants were using appropriate memories to complete the YAP (eg, recalling the previous week and not general routines).

Individual vs Group-Level Inference

Background

The sources of error described previously can have different implications on physical activity output obtained from self-report tools. Most likely, this output will be used to predict individual and/or group-level physical activity. The concept of estimation at the individual and at the group level can be confusing. It is tempting to think of a group as composed of individuals and, thus, to base inference for the group on individual-level inference. However, these indicators must be examined separately. Reliable estimation of usual physical activity at the individual level is more challenging and may require multiple assessments, but individual data can be used to obtain reliable and valid estimates of the distribution of usual physical activity in the group. Several studies have demonstrated how there is substantial variability in the agreement between individual estimates of self-report data and objective measures of physical activity [56-58]. Activity collected in a clinical setting (ie, for 1 individual alone) is a good example of the application of this level of measurement. On the other hand, if physical activity is collected in a group of individuals, the error is diluted across individuals and group-level physical activity scores can actually become close to "true" physical activity scores estimated for the group. In other words, the calibration design as described in this paper can minimize the distance of the individual

XSL•FO
RenderX

estimates to a line of best fit that represents the group (ie, line of best fit using least squares estimation). If standard regression assumptions hold, this procedure is known to create evenly distributed residuals that average out to a value of zero. Although individual error relies on the absolute distance between each individual observation and the line of best fit, the group estimate partials out the error to produce a close estimate of aggregate scores obtained from individuals.

YPAMS Example

In our development work with the YAP, we determined that individual error obtained from a calibration equation to predict minutes of MVPA from the PAQ can range from -56% to 69% of mean accelerometer values [58]. This wide range of agreement indicated that the PAQ tools might not be appropriate when used to quantify an individual's activity. However, we also found that group-level estimates ranged between -6% and 16% of average accelerometer scores [58]. This notion is important because it has important implications for survey research. The concept of group-level estimates and the potential of self-report tools for surveillance research might partially explain why self-report tools are still considered to be a valuable tool in physical activity research.

Statistical Analyses: Calibration

Background

Depending on the nature of physical activity questionnaires, it may be possible to generate individual calibration equations for discrete periods of time. An advantage of generating individual item regression calibration models is that context-adjusted beta weights (eg, β recess, β after school) can then be aggregated into a composite score (total weekly physical activity) by properly weighting the frequencies of each period. There are a number of other advantages of computing individual regression models for each item. In a more generic calibration approach [58], activity monitor data would not be fragmented into individual time segments. Raw item scores would be simply aggregated into a composite score and then calibrated against total MVPA measured by the accelerometer. This would cause individual items to be weighted equally in the prediction of total physical activity. The simplistic approach also inherently assumes that relationships with objective data would be similar (ie, that there would be equivalent slopes and intercepts across items). Individual equations provide a more robust calibration approach, but another key advantage is that the sample sizes can be maximized. This is directly related to the compliance criteria used to screen and process accelerometer data. A traditional calibration design would typically require participants to have worn the monitor for a complete day across an entire week whereas individual calibration would limit compliance to the availability of data for specific periods of the day. In other words, a participant that only wore the accelerometer during the evening would be excluded based on a traditional calibration approach, but with individually calibrated items his/her data would be included in the calibration of the evening item. By screening periods individually, it is possible to maximize sample size and improve statistical power. A final advantage is that it is possible to aggregate school (eg, recess, PE, commuting to school, lunch) and nonschool (eg, before school, after school, evening, Saturday, and Sunday) activity estimates to create separate composite activity scores for these time segments. This provides more utility for school leaders interested in quantifying physical activity during school. It also provides more educational value for children/parents and is more powerful for research applications.

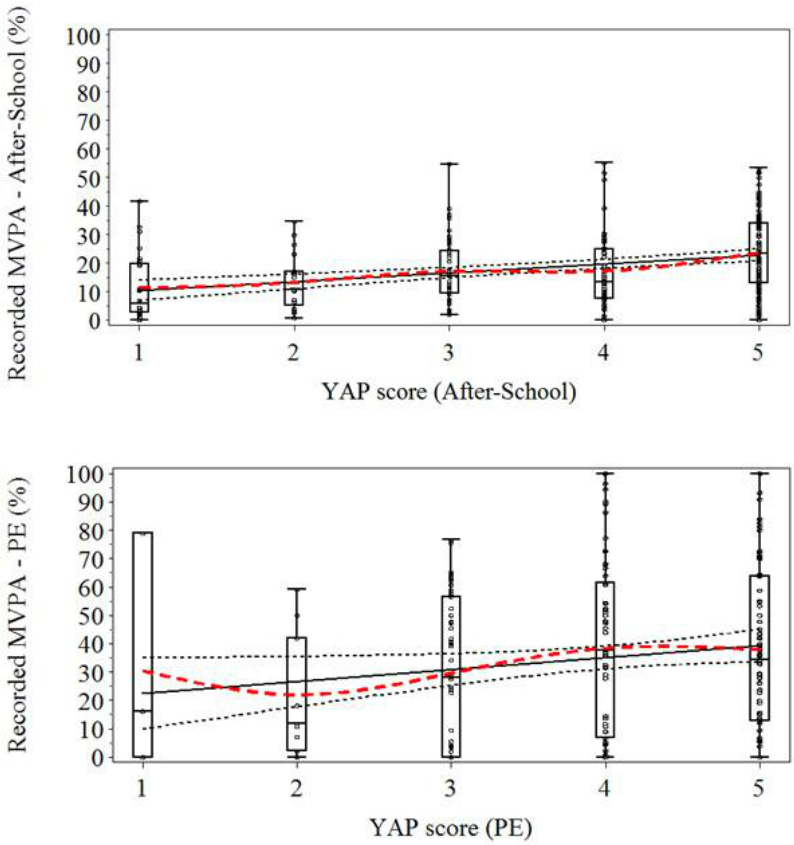
YPAMS Example

Our calibration work was specifically designed to calibrate individual items from the YAP against physical activity estimates obtained from the SenseWear Armband. An example of the calibration model for activity measured by the SenseWear Armband during PE using YAP item number 2 is provided by the equation $MVPA=\beta_0+\beta_1+\beta_2+\beta_3+\varepsilon_i$, in which *MVPA* is the SenseWear Armband measured percent time MVPA during PE class, β_0 is the intercept, β_1 is the beta weight associated with age, β_2 is the beta weight associated with gender, β_3 is the beta weight associated with YAP question 2 score (PE), and ε_i is random error (independent and normally distributed).

This approach is particularly important because our preliminary research has demonstrated that the assumption of equivalent slopes does not hold (unpublished results). Specifically, we found that some items are related linearly to accelerometer data whereas others are not. Figure 5 provides examples of the relation between the accelerometer and 2 YAP items as a result of our preliminary work. The bottom panel shows the relation between accelerometer EE values and YAP item 2 (activity during PE) whereas the top panel shows the relation between accelerometer EE values and YAP item 6 (activity after school). Observed PE estimates of activity have a fairly linear increase from a score of 2 through a score of 4 and a plateau at the highest end of the scale, whereas the same relation for activity after school is linear across the full-item scale. These relations can be accounted for by fitting different models for each of these items and, therefore, allowing regression coefficients to vary. By calibrating individual items, we have also maximized our sample size and found that using both approaches (individual calibration vs traditional week calibration protocols) on the same data resulted in a larger sample when items were individually calibrated (n=195 vs n=148, respectively). Finally, the individual YAP items are organized in 2 activity sections to provide separate estimates of activity at school and out-of-school contexts.

XSL•FO
RenderX

Figure 5. Relation between percent time in moderate-to-vigorous physical activity (MVPA) measured by the SenseWear Armband accelerometer and YAP raw scores during physical education (PE; bottom) and during after-school time (top) (n=221 participants from grades 4 to 12). The solid black line represents the line of best fit with respective 95% confidence intervals; the dashed red line fits a smooth curve across the distribution of scores. The lack of overlap between these suggests a nonlinear trend relation between percent time in MVPA and YAP scores.



Statistical Analyses: Cross-Validation

Background

Rather than describing the typical validation designs used in physical activity measurement research [44,56,59-64], we provide a brief explanation on the importance of alternative statistical designs to determine the accuracy of group-level estimates obtained from self-reports. Equivalence between physical activity measures is usually determined if the average difference between the 2 instruments is not statistically significant and, therefore, not different from zero. This approach has some limitations and, importantly, tests the hypothesis that measures obtained from the 2 instruments are different. This is the reverse of what is actually being tested. The classic approach for equivalence is also too conservative because it does not specify any margin of equivalence and therefore, depending on the design of the equivalence study (eg, large sample size), it might be oversensitive to small differences between self-report and accelerometer scores [65,66].

YPAMS Example

In the YPAMS, we evaluate group-level agreement using bioequivalence testing procedures. This method is commonly used in medical research and is often used to determine equivalency between alternate drug substances that have been developed for the same purpose [67]. For the purpose of calibration, we tested if the difference between accelerometer and self-report estimates of MVPA were within a predefined region of equivalence, defined as 10% (other equivalence regions can be defined, such as a 20% range). The region of equivalence is rather subjective, but there are specific recommendations proposed by the Food and Drug Administration [68] that can be used for guidance. Measures can be considered “equivalent” if the 2 1-sided 95% confidence intervals for the average difference between the 2 measures are within the equivalence region (ie, 10%) of the group accelerometer estimates.

Online Applications

Background

One key difference between print and online versions of a PAQ is that it is not clear if children will respond similarly to questions on a screen as they do with a print format. For example, the reliability and validity of Web-based surveys can be affected by the screen display of the survey (eg, color combinations, text background) [69,70]. The research on this topic is still limited, so it is premature to assume that youth would respond similarly to items on paper versus those on a screen. Therefore, the first step in this process is to directly evaluate possible differences between the print version and the new online version of the tool. Secondly, researchers also need to examine the feasibility of the online tool and anticipate administration, completion, and data access/sharing issues. One possible strategy to overcome unexpected challenges is to make

XSL•FO
RenderX

the survey available to a subset of participants/schools to simulate the broader use of the tool.

YPAMS Example

The YAP was developed with online applications in mind, but separate calibration work is needed because the Web format may alter relationships and associations. We have extended our calibration methods to enhance the utility of an online version and a game version deployed within the FITNESSGRAM software. In this case, it is important to refine the calibration equations while also testing for equivalence between formats. We conducted 2 small pilot studies to test these assumptions. In Study 1, we randomized 3 grade 5 classes (n=60) to complete the 3 different versions of the YAP 2 weeks apart. We found similar test-retest coefficients for print (r=.86), online (r=.79), and game (r=.84), which supported the utility of the Web-based versions. In Study 2, we recruited 70 children (aged 10-15 years) to test whether the existing print-based algorithms would work with the online version of the YAP. We found that the correlations between the measured and YAP physical activity indicators were high and significant (r=.70, P<.001); however, differences between measured and estimated minutes of physical activity were larger than in our previous work. This demonstrated the feasibility and potential of the online version and suggested that the online YAP can be improved if directly calibrated. These pilot studies demonstrate the importance of successive rounds of calibration and cross-validation to refine the accuracy and utility of the instruments over time. This principle would apply to any related calibration work with other instruments.

The preceding methodology has led to substantial reductions in overall error in our calibration models. Our preliminary models demonstrated that the individual calibration method resulted in a reduction of 44% of the error we would get if items were not individually calibrated. Although the results demonstrate good utility, our approach is to sequentially refine and improve the precision over time by increasing the size and diversity of the sample while also taking into account other factors, such as seasonality, urbanicity, and regional differences. These issues are introduced in the next section.

Additional Considerations

The inherent goal with any measurement instrument is to capture variation between individuals while minimizing sources of systematic and/or random error. Systematic error (bias) occurs consistently over measurements and may depend on subject-level attributes, whereas random error is the inconsistent variability among individuals or measurements. Unlike many other more stable health indicators, physical activity is a behavior that varies naturally from day to day and from season to season for an individual [71,72] and this makes it particularly challenging to assess. Assessing physical activity and sedentary behavior in youth is even more challenging due to maturation differences and sporadic activity patterns [73].

Considerable advances have been made in the use of objective tools (eg, accelerometers) in the past decade and there are similar opportunities to advance the quality and utility of self-report measures. The observed "error" in self-report instruments is often attributed to participant's bias or inability to recall information. These are certainly important factors, but it is also important to acknowledge that error is contributed directly by poorly designed questions, weak scoring methods, and an inability to accurately characterize and quantify the information that is provided. With calibration methods, it is possible to convert self-reported data into estimates that more closely model estimates that are obtained from objective sources. The information in this paper describes lessons learned through our exploration of ways to improve physical activity assessment protocols [44,73,74] and our interest in systematic efforts to improve the utility of self-report measures [13,22,58,75,76].

Another issue that needs to be tested is whether calibration would vary by region/season. Activity profiles vary by season and youth may respond differently about physical activity patterns in each season. Our design incorporates the natural variability across seasons and we are working to evaluate possible differences across regions. By replicating established calibration measurement protocols on independent samples of youth from different communities (at the same time), it will be possible to eventually determine whether calibration models hold when used in different regions. Cross-validation of the calibration equations originally developed would provide information about how the models hold in more diverse samples with different weather and culture. If there are differences, we anticipate being able to eventually develop more robust measurement models similar to those used in our adult calibration work [77].

A final need for continued development is to test the utility of the tool under less-controlled conditions. The original data for our calibration models were also collected by trained staff that could prompt youth to pay attention to aspects of the instrument. This helps to improve internal validity of the study, but may detract from external validity. It will eventually be important to test the accuracy of this tool under more real-world conditions. The distinction is similar to the way that preliminary calibration studies with accelerometers used controlled laboratory conditions and tested only locomotor activities. Subsequent studies conducted under free-living conditions demonstrated that it is considerably harder to accurately capture the diverse range of activities that children perform. Similarly, it will be more difficult to obtain accurate self-report data when used in less-controlled school settings. These points should not be interpreted as unsolvable problems, but as challenges to be overcome. We expect that calibration methods will enable accurate group-level estimates of physical activity to provide more accurate reports of age and gender patterns of physical activity; however, we acknowledge that it will remain difficult to obtain accurate individual estimates. Again, this is the same challenge faced by researchers working to refine the accuracy of objective measures. Calibration equations for monitors may have utility for group estimation, but accurately estimating individual data remains more elusive [56,57].

The refinement of prediction algorithms for the YAP has important school health and public health implications. Evaluating and refining the calibrations of the online and game-based versions of the YAP will facilitate planned adoption

XSL•FO
RenderX

within the FITNESSGRAM program. The newly established Presidential Youth Fitness Program (PYFP) recently established FITNESSGRAM as the national fitness test so this calibrated physical activity assessment tool would potentially provide a way to capture physical activity levels in schools throughout the country. Because calibration equations can be incorporated directly into the online tool, it would be possible to provide immediate feedback to youth and to generate automated school-level reports with aggregated data. Fully refined online versions of the YAP would enable broader use by schools interested in tracking participation in MVPA as part of district or state programming. The tools would also facilitate research applications on school activity (for epidemiology studies or behavioral interventions).

We anticipate opportunities to also explore potential for use in cross-cultural studies and international comparisons [78]. At this level, there will be language barriers and, expectedly, more cross-cultural differences that will need to be addressed. There are several procedures involved in this process and they target language barriers initially by using well-known translation procedures [79]. The International Physical Activity Questionnaire is a good example of the result of a combined effort to standardize and promote physical activity across the globe [80]. Several studies have used this tool and some examples of cross-validation studies include study populations of older adults in Japan [81], adults in Greenland [82], and adolescents in Vietnam [83]. We envision that similar developments and refinement are possible with the YAP.

Ultimately, we view this calibration work as a long-term, iterative process that will lead to continued and incremental improvements over time. The principles described in this paper utilize recommended practices to reduce measurement error with self-report measures [45] as well as recommended steps to test the validity of self-report instruments [46]. They provide a good guide to ensure that the work progresses in a systematic way. Although the calibration principles described here are specific for the YAP, the concepts and methods may have utility for researchers interested in similar calibration work with other tools or Web-based applications. There is an increased interest in Web applications of physical activity surveys [14-17,19,84]; however, to our knowledge no research has described a systematic process to develop, calibrate, and disseminate the use of such assessments.

Authors' Contributions

Both PSM and GW designed the YPAMS and have worked together on the draft of this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Youth Activity Profile Questionnaire (paper-version).

[PDF File (Adobe PDF File), 39KB - jmir_v16i12e269_app1.pdf]

References

1. 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]
2. Rowlands A. Accelerometer assessment of physical activity in children: an update. Pediatr Exerc Sci 2007 Aug;19(3):252-266. [Medline: 18019585]
3. Bauman A, Phongsavan P, Schoeppe S, Owen N. Physical activity measurement--a primer for health promotion. Promot Educ 2006;13(2):92-103. [Medline: 17017286]
4. Troiano RP. A timely meeting: objective measurement of physical activity. Med Sci Sports Exerc 2005 Nov;37(11 Suppl):S487-S489. [Medline: 16294111]
5. Trost SG, Ward DS, Moorehead SM, Watson PD, Riner W, Burke JR. Validity of the computer science and applications (CSA) activity monitor in children. Med Sci Sports Exerc 1998 Apr;30(4):629-633. [Medline: 9565947]
6. Eston RG, Rowlands AV, Ingledew DK. Validity of heart rate, pedometry, and accelerometry for predicting the energy cost of children's activities. J Appl Physiol (1985) 1998 Jan;84(1):362-371 [FREE Full text] [Medline: 9451658]
7. Puyau MR, Adolph AL, Vohra FA, Butte NF. Validation and calibration of physical activity monitors in children. Obes Res 2002 Mar;10(3):150-157. [doi: 10.1038/oby.2002.24] [Medline: 11886937]
8. Ekelund U, Sjöström M, Yngve A, Poortvliet E, Nilsson A, Froberg K, et al. Physical activity assessed by activity monitor and doubly labeled water in children. Med Sci Sports Exerc 2001 Feb;33(2):275-281. [Medline: 11224818]
9. Treuth MS, Schmitz K, Catellier DJ, McMurray RG, Murray DM, Almeida MJ, et al. Defining accelerometer thresholds for activity intensities in adolescent girls. Med Sci Sports Exerc 2004 Jul;36(7):1259-1266 [FREE Full text] [Medline: 15235335]
10. Mattocks C, Leary S, Ness A, Deere K, Saunders J, Tilling K, et al. Calibration of an accelerometer during free-living activities in children. Int J Pediatr Obes 2007;2(4):218-226. [doi: 10.1080/17477160701408809] [Medline: 17852552]

XSL•FO
RenderX

11. Schmitz KH, Treuth M, Hannan P, McMurray R, Ring KB, Catellier D, et al. Predicting energy expenditure from accelerometry counts in adolescent girls. Med Sci Sports Exerc 2005 Jan;37(1):155-161 [FREE Full text] [Medline: 15632682]
12. Bowles HR. Measurement of active and sedentary behaviors: closing the gaps in self-report methods. J Phys Act Health 2012 Jan;9 Suppl 1:S1-S4. [Medline: 22287442]
13. Troiano RP, Pettee Gabriel KK, Welk GJ, Owen N, Sternfeld B. Reported physical activity and sedentary behavior: why do you ask? J Phys Act Health 2012 Jan;9 Suppl 1:S68-S75. [Medline: 22287450]
14. Spook JE, Paulussen T, Kok G, Van Empelen P. Monitoring dietary intake and physical activity electronically: feasibility, usability, and ecological validity of a mobile-based Ecological Momentary Assessment tool. J Med Internet Res 2013;15(9):e214 [FREE Full text] [doi: 10.2196/jmir.2617] [Medline: 24067298]
15. Riebl SK, Paone AC, Hedrick VE, Zoellner JM, Estabrooks PA, Davy BM. The comparative validity of interactive multimedia questionnaires to paper-administered questionnaires for beverage intake and physical activity: pilot study. JMIR Res Protoc 2013;2(2):e40 [FREE Full text] [doi: 10.2196/resprot.2830] [Medline: 24148226]
16. da Costa FF, Schmoelz CP, Davies VF, Di Pietro PF, Kupek E, de Assis MA. Assessment of diet and physical activity of brazilian schoolchildren: usability testing of a web-based questionnaire. JMIR Res Protoc 2013;2(2):e31 [FREE Full text] [doi: 10.2196/resprot.2646] [Medline: 23958804]
17. Pearce PF, Williamson J, Harrell JS, Wildemuth BM, Solomon P. The children's computerized physical activity reporter: children as partners in the design and usability evaluation of an application for self-reporting physical activity. Comput Inform Nurs 2007;25(2):93-105. [doi: 10.1097/01.NCN.0000263979.54048.85] [Medline: 17356331]
18. Calabro MA, Welk GJ, Carriquiry AL, Nusser SM, Beyler NK, Mathews CE. Validation of a computerized 24-hour physical activity recall (24PAR) instrument with pattern-recognition activity monitors. J Phys Act Health 2009 Mar;6(2):211-220. [Medline: 19420399]
19. Namba H, Yamaguchi Y, Yamada Y, Tokushima S, Hatamoto Y, Sagayama H, et al. Validation of Web-based physical activity measurement systems using doubly labeled water. J Med Internet Res 2012;14(5):e123 [FREE Full text] [doi: 10.2196/jmir.2253] [Medline: 23010345]
20. Pfaeffli L, Maddison R, Jiang Y, Dalleck L, Löf M. Measuring physical activity in a cardiac rehabilitation population using a smartphone-based questionnaire. J Med Internet Res 2013;15(3):e61 [FREE Full text] [doi: 10.2196/jmir.2419] [Medline: 23524251]
21. Kipnis V, Carroll RJ, Freedman LS, Li L. Implications of a new dietary measurement error model for estimation of relative risk: application to four calibration studies. Am J Epidemiol 1999 Sep 15;150(6):642-651 [FREE Full text] [Medline: 10490004]
22. Nusser SM, Beyler NK, Welk GJ, Carriquiry AL, Fuller WA, King BMN. Modeling errors in physical activity recall data. J Phys Act Health 2012 Jan;9 Suppl 1:S56-S67. [Medline: 22287449]
23. Nusser S, Carriquiry A, Dodd K, Fuller W. A semiparametric transformation approach to estimating usual daily intake distributions. Journal of the American Statistical Association 1996;91:1440-1449.
24. Tooze JA, Troiano RP, Carroll RJ, Moshfegh AJ, Freedman LS. A measurement error model for physical activity level as measured by a questionnaire with application to the 1999-2006 NHANES questionnaire. Am J Epidemiol 2013 Jun 1;177(11):1199-1208 [FREE Full text] [doi: 10.1093/aje/kws379] [Medline: 23595007]
25. Neuhouser ML, Di C, Tinker LF, Thomson C, Sternfeld B, Mossavar-Rahmani Y, et al. Physical activity assessment: biomarkers and self-report of activity-related energy expenditure in the WHI. Am J Epidemiol 2013 Mar 15;177(6):576-585 [FREE Full text] [doi: 10.1093/aje/kws269] [Medline: 23436896]
26. Crocker PR, Bailey DA, Faulkner RA, Kowalski KC, McGrath R. Measuring general levels of physical activity: preliminary evidence for the Physical Activity Questionnaire for Older Children. Med Sci Sports Exerc 1997 Oct;29(10):1344-1349. [Medline: 9346166]
27. Kowalski K, Crocker PRE, Faulkner RA. Validation of the Physical Activity Questionnaire for Older Children. Pediatr Exerc Sci 1997;9:174-186.
28. Kowalski C, Crocker PRE, Kowalski NP. Convergent validity of the physical activity questionnaire for adolescents. Pediatr Exerc Sci 1997;9:342-352.
29. Bailey DA. The Saskatchewan Pediatric Bone Mineral Accrual Study: bone mineral acquisition during the growing years. Int J Sports Med 1997 Jul;18 Suppl 3:S191-S194. [doi: 10.1055/s-2007-972713] [Medline: 9272847]
30. Crocker PR, Eklund RC, Kowalski KC. Children's physical activity and physical self-perceptions. J Sports Sci 2000 Jun;18(6):383-394. [doi: 10.1080/02640410050074313] [Medline: 10902673]
31. Moore JB, Hanes JC, Barbeau P, Gutin B, Treviño RP, Yin Z. Validation of the Physical Activity Questionnaire for Older Children in children of different races. Pediatr Exerc Sci 2007 Feb;19(1):6-19. [Medline: 17554153]
32. Janz KF, Lutuchy EM, Wenthe P, Levy SM. Measuring activity in children and adolescents using self-report: PAQ-C and PAQ-A. Med Sci Sports Exerc 2008 Apr;40(4):767-772. [doi: 10.1249/MSS.0b013e3181620ed1] [Medline: 18317366]
33. Martínez-Gómez D, Martínez-de-Haro V, Pozo T, Welk GJ, Villagra A, Calle ME, et al. Reliability and validity of the PAQ-A questionnaire to assess physical activity in Spanish adolescents. Revista Espanola de Salud Publica 2009;83:427-439.

XSL•FO
RenderX

34. Carroll R, Ruppert D, Stefanski L, Crainiceanu C. Regression calibration. In: Measurement Error in Nonlinear Models: A Modern Perspective 2nd edition. Boca Raton, FL: Taylor and Francis Group; 2006:65-95.
35. Cradock AL, Barrett JL, Carter J, McHugh A, Sproul J, Russo ET, et al. Impact of the Boston Active School Day policy to promote physical activity among children. Am J Health Promot 2014;28(3 Suppl):S54-S64. [doi: 10.4278/ajhp.130430-QUAN-204] [Medline: 24380467]
36. Mendoza JA, Watson K, Nguyen N, Cerin E, Baranowski T, Nicklas TA. Active commuting to school and association with physical activity and adiposity among US youth. J Phys Act Health 2011 May;8(4):488-495 [FREE Full text] [Medline: 21597121]
37. Ruch N, Joss F, Jimmy G, Melzer K, Hänggi J, Mäder U. Neural network versus activity-specific prediction equations for energy expenditure estimation in children. J Appl Physiol (1985) 2013 Nov 1;115(9):1229-1236 [FREE Full text] [doi: 10.1152/japplphysiol.01443.2012] [Medline: 23990244]
38. Pober DM, Staudenmayer J, Raphael C, Freedson PS. Development of novel techniques to classify physical activity mode using accelerometers. Med Sci Sports Exerc 2006 Sep;38(9):1626-1634. [doi: 10.1249/01.mss.0000227542.43669.45] [Medline: 16960524]
39. Choi L, Liu Z, Matthews CE, Buchowski MS. Validation of accelerometer wear and nonwear time classification algorithm. Med Sci Sports Exerc 2011 Feb;43(2):357-364 [FREE Full text] [doi: 10.1249/MSS.0b013e3181ed61a3] [Medline: 20581716]
40. Lee IM, Shiroma EJ. Using accelerometers to measure physical activity in large-scale epidemiological studies: issues and challenges. Br J Sports Med 2014 Feb;48(3):197-201. [doi: 10.1136/bjsports-2013-093154] [Medline: 24297837]
41. Arvidsson D, Slinde F, Larsson S, Hulthén L. Energy cost of physical activities in children: validation of SenseWear Armband. Med Sci Sports Exerc 2007 Nov;39(11):2076-2084. [doi: 10.1249/mss.0b013e31814fb439] [Medline: 17986918]
42. Jakicic JM, Marcus M, Gallagher KI, Randall C, Thomas E, Goss FL, et al. Evaluation of the SenseWear Pro Armband to assess energy expenditure during exercise. Med Sci Sports Exerc 2004 May;36(5):897-904. [Medline: 15126727]
43. Bassett DR, Rowlands A, Trost SG. Calibration and validation of wearable monitors. Med Sci Sports Exerc 2012 Jan;44(1 Suppl 1):S32-S38 [FREE Full text] [doi: 10.1249/MSS.0b013e3182399cf7] [Medline: 22157772]
44. Welk GJ. Principles of design and analyses for the calibration of accelerometry-based activity monitors. Med Sci Sports Exerc 2005 Nov;37(11 Suppl):S501-S511. [Medline: 16294113]
45. Gracia-Marco L, Ortega FB, Ruiz JR, Williams CA, Hagströmer M, Manios Y, Helena Study Group. Seasonal variation in physical activity and sedentary time in different European regions. The HELENA study. J Sports Sci 2013;31(16):1831-1840. [doi: 10.1080/02640414.2013.803595] [Medline: 24050788]
46. Whitlock G, Clark T, Vander Hoorn S, Rodgers A, Jackson R, Norton R, et al. Random errors in the measurement of 10 cardiovascular risk factors. Eur J Epidemiol 2001;17(10):907-909. [Medline: 12188008]
47. Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu CM. Linear regression and atenuation. In: Measurement Error in Nonlinear Models: A Modern Perspective. Volume 2. Boca Raton, FL: Chapman & Hall/CRC; 2006:41-64.
48. Ferrari P, Friedenreich C, Matthews CE. The role of measurement error in estimating levels of physical activity. Am J Epidemiol 2007 Oct 1;166(7):832-840 [FREE Full text] [doi: 10.1093/aje/kwm148] [Medline: 17670910]
49. Harrell JS, McMurray RG, Baggett CD, Pennell ML, Pearce PF, Bangdiwala SI. Energy costs of physical activities in children and adolescents. Med Sci Sports Exerc 2005 Feb;37(2):329-336. [Medline: 15692331]
50. Ridley K, Olds TS. Assigning energy costs to activities in children: a review and synthesis. Med Sci Sports Exerc 2008 Aug;40(8):1439-1446. [doi: 10.1249/MSS.0b013e31817279ef] [Medline: 18614948]
51. Esliger DW, Copeland JL, Barnes JD, Tremblay MS. Standardizing and optimizing the use of accelerometer data for free-living physical activity monitoring. J Phys Act Health 2005;3:366-383.
52. De Meester F, De Bourdeaudhuij I, Deforche B, Ottevaere C, Cardon G. Measuring physical activity using accelerometry in 13-15-year-old adolescents: the importance of including non-wear activities. Public Health Nutr 2011 Dec;14(12):2124-2133. [doi: 10.1017/S1368980011001868] [Medline: 21835080]
53. Ottevaere C, Huybrechts I, De Meester F, De Bourdeaudhuij I, Cuenca-Garcia M, De Henauw S. The use of accelerometry in adolescents and its implementation with non-wear time activity diaries in free-living conditions. J Sports Sci 2011 Jan;29(1):103-113. [doi: 10.1080/02640414.2010.521169] [Medline: 21104522]
54. Mâsse LC, Fuemmeler BF, Anderson CB, Matthews CE, Trost SG, Catellier DJ, et al. Accelerometer data reduction: a comparison of four reduction algorithms on select outcome variables. Med Sci Sports Exerc 2005 Nov;37(11 Suppl):S544-S554. [Medline: 16294117]
55. Toftager M, Kristensen PL, Oliver M, Duncan S, Christiansen LB, Boyle E, et al. Accelerometer data reduction in adolescents: effects on sample retention and bias. Int J Behav Nutr Phys Act 2013;10:140 [FREE Full text] [doi: 10.1186/1479-5868-10-140] [Medline: 24359480]
56. Hallal PC, Reichert FF, Clark VL, Cordeira KL, Menezes AM, Eaton S, et al. Energy expenditure compared to physical activity measured by accelerometry and self-report in adolescents: a validation study. PLoS One 2013;8(11):e77036 [FREE Full text] [doi: 10.1371/journal.pone.0077036] [Medline: 24223707]

XSL•FO
RenderX

57. Corder K, van Sluijs EM, Wright A, Whincup P, Wareham NJ, Ekelund U. Is it possible to assess free-living physical activity and energy expenditure in young people by self-report? Am J Clin Nutr 2009 Mar;89(3):862-870 [FREE Full text] [doi: 10.3945/ajcn.2008.26739] [Medline: 19144732]
58. Saint-Maurice PF, Welk GJ, Beyler NK, Bartee RT, Heelan KA. Calibration of self-report tools for physical activity research: the Physical Activity Questionnaire (PAQ). BMC Public Health 2014;14:461 [FREE Full text] [doi: 10.1186/1471-2458-14-461] [Medline: 24886625]
59. Scholes S, Coombs N, Pedisic Z, Mindell JS, Bauman A, Rowlands AV, et al. Age- and sex-specific criterion validity of the health survey for England Physical Activity and Sedentary Behavior Assessment Questionnaire as compared with accelerometry. Am J Epidemiol 2014 Jun 15;179(12):1493-1502 [FREE Full text] [doi: 10.1093/aje/kwu087] [Medline: 24863551]
60. Wang C, Chen P, Zhuang J. Validity and reliability of International Physical Activity Questionnaire-Short Form in Chinese youth. Res Q Exerc Sport 2013 Dec;84 Suppl 2:S80-S86. [doi: 10.1080/02701367.2013.850991] [Medline: 24527570]
61. Medina C, Barquera S, Janssen I. Validity and reliability of the International Physical Activity Questionnaire among adults in Mexico. Rev Panam Salud Publica 2013 Jul;34(1):21-28 [FREE Full text] [Medline: 24006016]
62. Manios Y, Androutsos O, Moschonis G, Birbilis M, Maragkopoulou K, Giannopoulou A, et al. Criterion validity of the Physical Activity Questionnaire for Schoolchildren (PAQ-S) in assessing physical activity levels: the Healthy Growth Study. J Sports Med Phys Fitness 2013 Oct;53(5):502-508. [Medline: 23903530]
63. Matthews CE, Keadle SK, Sampson J, Lyden K, Bowles HR, Moore SC, et al. Validation of a previous-day recall measure of active and sedentary behaviors. Med Sci Sports Exerc 2013 Aug;45(8):1629-1638 [FREE Full text] [doi: 10.1249/MSS.0b013e3182897690] [Medline: 23863547]
64. Dyrstad SM, Hansen BH, Holme IM, Anderssen SA. Comparison of self-reported versus accelerometer-measured physical activity. Med Sci Sports Exerc 2014 Jan;46(1):99-106. [doi: 10.1249/MSS.0b013e3182a0595f] [Medline: 23793232]
65. Walker E, Nowacki AS. Understanding equivalence and noninferiority testing. J Gen Intern Med 2011 Feb;26(2):192-196 [FREE Full text] [doi: 10.1007/s11606-010-1513-8] [Medline: 20857339]
66. Robinson AP, Froese RE. Model validation using equivalence tests. Ecological Modelling 2004 Sep;176(3-4):349-358. [doi: 10.1016/j.ecolmodel.2004.01.013]
67. Schuirmann DJ. A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. J Pharmacokinet Biopharm 1987 Dec;15(6):657-680. [Medline: 3450848]
68. Midha KK, McKay G. Bioequivalence; its history, practice, and future. AAPS J 2009 Dec;11(4):664-670 [FREE Full text] [doi: 10.1208/s12248-009-9142-z] [Medline: 19806461]
69. Hall R, Hanna P. The impact of web page text-background colour combinations on readability, retention, aesthetics and behavioural intention. Behaviour & Information Technology 2004;23:183-195.
70. Gnambs T, Appel M, Batinic B. Color red in web-based knowledge testing. Computers in Human Behavior 2010;26:1625-1631.
71. Baranowski T, Thompson WO, DuRant RH, Baranowski J, Puhl J. Observations on physical activity in physical locations: age, gender, ethnicity, and month effects. Res Q Exerc Sport 1993 Jun;64(2):127-133. [doi: 10.1080/02701367.1993.10608789] [Medline: 8341835]
72. Tucker P, Gilliland J. The effect of season and weather on physical activity: a systematic review. Public Health 2007 Dec;121(12):909-922. [doi: 10.1016/j.puhe.2007.04.009] [Medline: 17920646]
73. Welk GJ, Corbin CB, Dale D. Measurement issues in the assessment of physical activity in children. Res Q Exerc Sport 2000 Jun;71(2 Suppl):S59-S73. [Medline: 10925827]
74. Welk GJ, McClain J, Ainsworth BE. Protocols for evaluating equivalency of accelerometry-based activity monitors. Med Sci Sports Exerc 2012 Jan;44(1 Suppl 1):S39-S49. [doi: 10.1249/MSS.0b013e3182399d8f] [Medline: 22157773]
75. Welk GJ, Kim Y, Stanfill B, Osthus DA, Calabro MA, Nusser SM, et al. Validity of 24-h physical activity recall: physical activity measurement survey. Med Sci Sports Exerc 2014 Oct;46(10):2014-2024. [doi: 10.1249/MSS.0000000000000314] [Medline: 24561818]
76. Tucker JM, Welk G, Nusser SM, Beyler NK, Dzewaltowski D. Estimating minutes of physical activity from the previous day physical activity recall: validation of a prediction equation. J Phys Act Health 2011 Jan;8(1):71-78. [Medline: 21297187]
77. Osthus D, Beyler NK, Stanfill B, Nusser SM, Fuller WA, Carriquiry AL, et al. Estimating the distribution of usual daily energy expenditure for subpopulations from a probability sample. Statistics in Medicine 2014:1-16.
78. Hallal PC, Andersen LB, Bull FC, Guthold R, Haskell W, Ekelund U, Lancet Physical Activity Series Working Group. Global physical activity levels: surveillance progress, pitfalls, and prospects. Lancet 2012 Jul 21;380(9838):247-257. [doi: 10.1016/S0140-6736(12)60646-1] [Medline: 22818937]
79. Sousa VD, Rojjanasrirat W. Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user-friendly guideline. J Eval Clin Pract 2011 Apr;17(2):268-274. [doi: 10.1111/j.1365-2753.2010.01434.x] [Medline: 20874835]
80. Bull FC, Maslin TS, Armstrong T. Global physical activity questionnaire (GPAQ): nine country reliability and validity study. J Phys Act Health 2009 Nov;6(6):790-804. [Medline: 20101923]

XSL•FO
RenderX

81. Tomioka K, Iwamoto J, Saeki K, Okamoto N. Reliability and validity of the International Physical Activity Questionnaire (IPAQ) in elderly adults: the Fujiwara-kyo Study. J Epidemiol 2011;21(6):459-465 [FREE Full text] [Medline: 21946625]
82. Dahl-Petersen IK, Hansen AW, Bjerregaard P, Jørgensen ME, Brage S. Validity of the international physical activity questionnaire in the arctic. Med Sci Sports Exerc 2013 Apr;45(4):728-736. [doi: 10.1249/MSS.0b013e31827a6b40] [Medline: 23190587]
83. Lachat CK, Verstraeten R, Khanh le NB, Hagströmer M, Khan NC, Van Ndo A, et al. Validity of two physical activity questionnaires (IPAQ and PAQA) for Vietnamese adolescents in rural and urban areas. Int J Behav Nutr Phys Act 2008;5:37 [FREE Full text] [doi: 10.1186/1479-5868-5-37] [Medline: 18616798]
84. De Vera MA, Ratzlaff C, Doerfling P, Kopec J. Reliability and validity of an internet-based questionnaire measuring lifetime physical activity. Am J Epidemiol 2010 Nov 15;172(10):1190-1198 [FREE Full text] [doi: 10.1093/aje/kwq273] [Medline: 20876666]

Abbreviations

EE: energy expenditure
MET: metabolic equivalent of task
MVPA: moderate-to-vigorous physical activity
PAQ: Physical Activity Questionnaire
PE: physical education
PYFP: Presidential Youth Fitness Program
YAP: Youth Activity Profile
YPAMS: Youth Physical Activity Measurement Study

Edited by G Eysenbach; submitted 02.07.14; peer-reviewed by A Rebar, L Ning; comments to author 25.07.14; revised version received 04.09.14; accepted 16.09.14; published 01.12.14.

Please cite as:
Saint-Maurice PF, Welk GJ
Web-Based Assessments of Physical Activity in Youth: Considerations for Design and Scale Calibration
J Med Internet Res 2014;16(12):e269
URL: http://www.jmir.org/2014/12/e269/
doi:10.2196/jmir.3626
PMID:25448192

©Pedro F Saint-Maurice, Gregory J Welk. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 01.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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

Interformat Reliability of Digital Psychiatric Self-Report Questionnaires: A Systematic Review

Sven Alfonsson[1*], PhD; Pernilla Maathz[2*], MSc; Timo Hursti[3], PhD

[1]U-CARE, Department of Public Health and Caring Sciences, Uppsala University, Uppsala, Sweden

[2]U-CARE, Department of Psychology, Uppsala University, Uppsala, Sweden

[3]Department of Psychology, Uppsala University, Uppsala, Sweden

[*]these authors contributed equally

Corresponding Author:
Sven Alfonsson, PhD
U-CARE
Department of Public Health and Caring Sciences
Uppsala University
A11
Dag Hammarskjölds väg BMC
Uppsala, 751 22
Sweden
Phone: 46 0184716194
Fax: 46 0184716675
Email: sven.alfonsson@pubcare.uu.se

Abstract

Background: Research on Internet-based interventions typically use digital versions of pen and paper self-report symptom scales. However, adaptation into the digital format could affect the psychometric properties of established self-report scales. Several studies have investigated differences between digital and pen and paper versions of instruments, but no systematic review of the results has yet been done.

Objective: This review aims to assess the interformat reliability of self-report symptom scales used in digital or online psychotherapy research.

Methods: Three databases (MEDLINE, Embase, and PsycINFO) were systematically reviewed for studies investigating the reliability between digital and pen and paper versions of psychiatric symptom scales.

Results: From a total of 1504 publications, 33 were included in the review, and interformat reliability of 40 different symptom scales was assessed. Significant differences in mean total scores between formats were found in 10 of 62 analyses. These differences were found in just a few studies, which indicates that the results were due to study effects and sample effects rather than unreliable instruments. The interformat reliability ranged from r=.35 to r=.99; however, the majority of instruments showed a strong correlation between format scores. The quality of the included studies varied, and several studies had insufficient power to detect small differences between formats.

Conclusions: When digital versions of self-report symptom scales are compared to pen and paper versions, most scales show high interformat reliability. This supports the reliability of results obtained in psychotherapy research on the Internet and the comparability of the results to traditional psychotherapy research. There are, however, some instruments that consistently show low interformat reliability, suggesting that these conclusions cannot be generalized to all questionnaires. Most studies had at least some methodological issues with insufficient statistical power being the most common issue. Future studies should preferably provide information about the transformation of the instrument into digital format and the procedure for data collection in more detail.

(J Med Internet Res 2014;16(12):e268) doi:10.2196/jmir.3395

KEYWORDS

psychometric; reliability; questionnaire; psychotherapy; computer; Internet

XSL•FO
RenderX

Introduction

The use of computerized psychological assessment has increased, and today many patients are also helped through Internet-based psychological interventions. The efficacy of Internet-based interventions has been evaluated repeatedly, and the effects seem to be comparable to live interventions [1-3]. When collecting psychiatric data through computers or evaluating Internet interventions, researchers typically rely on digital versions of existing pen and paper (PnP) self-report instruments. However, it cannot be assumed that instruments retain their psychometric properties when the format of delivery is changed [4]. The level of equality between different delivery formats is here referred to as *interformat reliability*. A high interformat reliability indicates that the psychometric properties of the instrument are independent of the delivery format.

Interformat reliability could be affected in two main ways: by characteristics of the delivery format itself or by how respondents perceive the delivery format. Digital instruments can be presented on different platforms, for example, on a standalone computer, an online Web page, or a mobile phone. Each platform has its own interface, and scores could be affected as a result. The presentation of an instrument can also be different in other ways, for example, by presenting one item at a time as opposed to several items on the same page. Effects of interface and presentation of digital instruments have not been investigated empirically to any large extent, making the strength of such possible effects unclear [5,6]. Since the effects are uncertain, design choices regarding the adaptation of instruments to digital format are important. For example, it can be argued that differences in layout adaptation may affect the validity of the results [4].

Further, people might respond differently to a digital instrument depending on how they perceive the level of security and anonymity [7]. Some people may also feel uncomfortable in using digital devices, which may affect the results [8]. People may express themselves differently in a digital context, for example, when communicating via the Internet, as compared to face-to-face interactions [8,9]. For example, some very sensitive data may benefit from digital assessment [10,11]. If respondents' ratings on items and the resulting score are affected by presentation format, this could affect the conclusions drawn from Internet-based psychotherapy research.

Thus, investigations of the psychometric equivalence of computer- and PnP-administered instruments are warranted. Quite a few examinations of the interformat reliability of self-report questionnaires have been done, especially among somatic patients, but no systematic review focusing on psychiatric instruments has yet been conducted [12]. Such a review would be valuable in deciding whether transformation of questionnaires to online use are feasible and whether scores from pen and paper and digital versions can be compared.

The objective of this study was to review the interformat reliability of self-report symptom scales used in psychotherapy research. This review also aimed to assess the methodological quality of studies investigating interformat reliability.

Methods

Search Strategy

The review process was guided by the Cochrane Handbook of Interventions and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. Unfortunately, there are no specific guidelines for reviewing clinical measurement tools, such as self-report instruments, and the recommendations for judging risk of bias cannot be applied directly. As a consequence, a protocol for quality assessment was created for this review, based on the guidelines in the Cochrane handbook. A systematic search of the literature was performed in the databases MEDLINE, Embase, and PsycINFO. The search strategy included four concepts: the digital format, self-report questionnaires, psychometric properties, and psychology. Several search terms were used for each concept; for example, the search line in MEDLINE was "computer OR Internet OR online AND questionnaire OR instrument OR scale AND psychometric OR reliability OR validity AND psychology OR psychotherapy OR psychological". The reference lists of included publications were examined in order to identify additional relevant studies. The risk of bias may be high in such perusal of reference lists and is often discouraged. In the present case, the risk of bias was judged somewhat lower and counterbalanced against the benefit of finding older studies and papers published outside the usual channels. No attempt was made to locate unpublished material. The literature search was conducted between January and May 2013, and to identify any studies published on a later date, an additional search was done in January 2014.

Study Selection

All published peer-reviewed English language research studies comparing the psychometric properties of computerized and PnP versions of self-report instruments were considered for inclusion in the review. Study subjects had to be adults and data on interformat reliability reported as part of the results, either as a correlation, an analysis of differences between format mean scores, or as comparison between theoretical models.

Studies investigating instruments designed to measure symptoms of any of the following Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) Axis I group of diagnoses were included: mood disorders, anxiety disorders, eating disorders, substance use disorders, and sleep disorders. Instruments assessing personality traits or non-clinical behaviors (eg, exercise) were excluded. Only instruments of the questionnaire type with set answers and ratings were included; for example, not behavioral assessments, Visual Analogue Scale (VAS), diaries, or open questions. Also, only studies of instruments that previously had been psychometrically assessed in pen and paper format were included.

After the initial search, the title and abstract of identified publications were examined by the first and second authors independently. Irrelevant publications were subsequently excluded. All publications judged relevant were retrieved in full text. The first and second author independently reviewed all full text publications based on the inclusion and exclusion criteria. In case of disagreements, consensus was sought through

XSL•FO
RenderX

discussion. If agreement could not be reached, the third author's judgment was final.

Data Extraction Procedure

A systematic approach to data extraction was used to produce a descriptive summary of the methods used and the psychometric findings in each included study. The study characteristics (year of publication, sample size, administration format [computer, online, or palm/cell phone], design) and participant characteristics (population, age, gender, computer experience) were extracted. Psychometric data regarding each instrument's interformat reliability, as well as regarding test-retest reliability and internal consistency of the digital format, were extracted from each study. The first and second author independently extracted data from all included studies.

Quality Assessment

There are no established guidelines for assessing the quality of psychometric studies, and so a strategy was created for this study. The quality of each study was assessed and rated on six aspects: (1) type of analysis used to compare instruments, (2) use of a randomization procedure, (3) reporting of statistics and results, (4) sample size, (5) sample type, and (6) description of digital instrument adaptation and design. Each aspect was rated on a 3-point scale (0-2), providing a quality score with a possible range from 0-12. See Multimedia Appendix 1 for detailed descriptions of quality assessment. The first and second authors assessed each study independently and agreed on 90.3% of the quality assessment elements before discussion. Any remaining disagreement was discussed with the last author whose judgment was decisive. Studies obtaining a total score above two thirds of the maximum score (ie, >8) were considered high quality studies.

Study Designs

Interformat reliability can be investigated using either a one sample design or two samples design. Both designs can be further enhanced by adding randomization and crossover design elements. In one sample designs, a single sample is drawn from a population. Data for different formats of an instrument can then be collected by randomizing participants to either complete a PnP or a digital version of the instrument, or by letting each participant complete the instrument in both formats, that is, a crossover design. In the simpler forms of one sample design, participants either first complete one format and then the other without randomizing the order of formats, or are allocated to a group that completes only one of the two formats. These designs have major weaknesses since order effects cannot be separated, and there may be group differences. Hence, there is a clear risk of bias and randomized crossover design is to be preferred. Crossover design has the additional advantage of providing greater statistical power and thus requiring a smaller sample size. In a two samples design, two samples are drawn from one or two different populations. Participants in one of the samples complete the instrument in the PnP version, while participants in the other sample complete the instrument in the digital version. Since mean scores and variance cannot be assumed to be equal in two samples or populations, any conclusions about interformat reliability drawn from a study using a two sample design will be without scientific value.

An additional way to assess interformat reliability is to investigate whether the statistical model of an instrument is equal in both formats. When an instrument is designed, item- or factor analysis is often conducted. If the structure of the instrument can be replicated in the digital version, this will provide some evidence of interformat reliability. However, the structure can be equal in both formats while the actual scores on the instrument diverge, and therefore this should be used only as an addition to other analyses of equivalence.

Assessing Interformat Reliability

To be able to compare results from instruments in digital format with results from PnP format, one must investigate either correlations between scores or differences between mean scores. Reliability is typically measured with correlation analysis where a result closer to 1 is a stronger relationship [13]. While Pearson correlation is sometimes used when analyzing reliability, for interformat reliability intraclass correlation is more adequate. It is important to note that a high correlation between two scores does not imply that the scores are at the same level. For example, one score could be systematically lower than the other and the correlation would still be high. Hence, equality between different formats is better assessed with a statistical test of differences between group mean scores, for example, by *t* tests or analysis of variance (ANOVA). A study comparing instruments in two different formats should thus report both correlations and analysis of differences between mean scores, with the latter being essential.

Other forms of reliability commonly investigated in instruments are test-retest and internal reliability [13]. As with interformat reliability, it cannot be assumed that test-retest or internal reliability of the instrument is sustained when the format of delivery is changed.

Sample Size and Effect Size

Sample size calculations should be based on analysis of mean differences, something that typically requires larger sample sizes than correlation analysis. When calculating power and sample size for a study, one should decide how large differences between formats one is ready to accept. While many studies of Internet-based psychological interventions show medium to large effect sizes, it could be argued that when it comes to differences between instrument formats, even small effect sizes are important to detect. For correlation analysis, sample size calculations should be based on achieving adequate confidence intervals rather than significance testing [14].

Statistical Analyses

When statistical analyses of mean differences were not reported in the original publication, the authors performed the corresponding *t* test if required data were available. The authors also calculated effects sizes for differences where these were not reported in the original publications. Differences between groups of studies (ie, high and low quality studies) were analyzed with *t* tests or Mann-Whitney tests. Interformat reliability on an aggregate level was investigated by comparing

XSL•FO
RenderX

mean scores and using calculations of binominal probability. Cohen's *d* was used as a measure of effect size where .2 is considered a small effect, .5 a medium effect, and .8 a large effect size. A *P* value of <.05 was used as a threshold of statistical significance.

Results

Identified and Included Publications

The initial search yielded a total of 1504 hits in the databases. After review of titles and abstracts, 61 publications were selected for full text review. Of these 61 publications, 29 met inclusion criteria. An additional 8 publications were included after the examination of reference lists. Following the additional search, one more publication was included. In total, 38 publications were thus included for data extraction. A complete list of included studies [15-52] can be found in Table 1. A PRISMA flow diagram of the search and inclusion process can be found in Multimedia Appendix 2.

Data Extraction

Before discussion, reviewers agreed on 92.6% of extracted data elements regarding the study characteristics, 90.3% of the quality assessment elements, and 99.3% of the psychometric data. This is considered a high level of agreement. When studies did not report analysis of mean differences between formats, *t* tests were calculated by the reviewers to complete the results. When crossover studies did not report total format mean scores, this was calculated by the reviewers, based on the reported group mean scores. When two sample designs are used, equality between mean scores of the instruments cannot be assumed. Therefore, the five studies using this design were excluded for further analyses of interformat reliability but are included in Table 1 for the benefit of readers.

Study Characteristics

The included 33 publications were published between 1985 and 2013. All publications described unique studies, and most studies investigated more than one instrument. Sample sizes ranged from 29 to 1171 (mean 224, SD 277.5). A third of studies (11/33, 33%) assessed stand-alone computer administration, more than half (19/33, 58%) assessed online administration, and very few assessed palm device administration (2/33, 6%) and both online and smartphone administration (1/33, 3%). Nearly a quarter of the studies (8/33, 24%) included some assessment of computer experience. The included studies investigated 40 different self-report instruments, covering the following diagnoses or problem areas: Panic disorder, Depression, Anxiety, Eating disorders, Alcohol and tobacco dependence or misuse, Obsessive Compulsive Disorder, Posttraumatic stress, Postpartum Depression, Social Anxiety Disorder, Insomnia, and perceived physical and mental health.

Participant Characteristics

We found that 42% (14/33) of studies used a sample from patients or other appropriate population, 45% (15/33) used a student sample, 9% (3/33) used some kind of community sample, and 3% (1/33) did not define the sample. The mean age of the participants ranged between 18.8 and 68.3 years. The gender proportions of the samples ranged from 23.9% to 79.9% women in the studies that included both sexes. Two studies investigated screening instruments of postpartum depression and used all women samples. See Multimedia Appendix 3 for a complete list of study and participant characteristics.

Design and Quality of Included Studies

Of the 33 studies, 17 (52%) employed a crossover design. Most studies (29/33, 88%) reported adequate statistics, while more than half (17/33, 52%) did not describe the adaptation of the instruments to a digital format. The mean quality score was 8.6. Using a cut score of >8 (two thirds of total quality score), 20 of the 33 studies (61%) were assessed as high quality studies. See Multimedia Appendix 1 for complete quality assessment scores.

XSL•FO
RenderX

Table 1. Included publications, investigated instruments, formats compared, and study designs.

Publication	Instruments	Formats	Design
Austin et al (2006) [15]	Body Sensations Questionnaire, Agoraphobic Cognitions Questionnaire, Mobile Inventory	PnP, online	One sample crossover (2x2)
Brock et al (2012) [16]	Center for Epidemiological Studies-Depression, Beck Anxiety Inventory	PnP, online	One sample crossover (4x2)
Bush et al (2013) [17]	PTSD Check List–Civilian Version, Patient Health Questionnaire-9	PnP, online, smartphone	One sample randomized
Butler et al (1988) [18]	Setting Conditions for Anorexia Nervosa Scale	PnP, computer	One sample crossover (2x2)
Carlbring et al (2007) [19]	Body Sensations Questionnaire, Agoraphobic Cognitions Questionnaire, Mobile Inventory, Beck Anxiety Inventory, Beck Depression Inventory-II, Montgomery-Asberg Depression Rating Scale-Self-report	PnP, online	One sample crossover (2x2)
Chan-Pensley (1999) [20]	Alcohol Use Disorder Identification Test	PnP, computer	One sample crossover (2x2)
Coles et al (2007) [21]	Obsessive-Compulsive Inventory, Obsessive Beliefs Questionnaire-44	PnP, online	One sample crossover (2x2)
Cook et al (2007) [22]	Quick Inventory of Depressive Symptomatology Self Rated	PnP, palm	One sample crossover (2x2)
Fortson et al (2006) [23]	Center for Epidemiologic Studies Depression scale, Trauma Symptom Screen	PnP, online	One sample crossover (4x2)
George et al (1992) [24]	Beck Depression Inventory, State-Trait Anxiety Inventory-State, State-Trait Anxiety Inventory-Trait	PnP, computer	One sample randomized
Glaze & Cox (1991) [25]	Edinburgh Postnatal Depression Scale	PnP, computer	One sample crossover (2x2)
Herrero & Meneses (2006) [26]	Center for Epidemiologic Studies Depression scale -7	PnP, online	One sample randomized
Hirai et al (2011) [27]	Social Interaction Anxiety Scale, Social Phobia Scale	PnP, online	One sample randomized
Holländare et al (2008) [28]	Beck Depression Inventory-II, Montgomery-Asberg Depression Rating Scale-Self-report	PnP, online	One sample crossover (2x2)
Holländare et al (2010) [29]	Beck Depression Inventory-II, Montgomery-Asberg Depression Rating Scale-Self-report	PnP, online	One sample crossover (2x2)
Kurt et al (2004) [30]	Center for Epidemiologic Studies Depression scale -R-20, Geriatric Depression Scale-15	PnP, computer	One sample crossover (2x2)
Lankford et al (1994) [31]	Beck Depression Inventory, State-Trait Anxiety Inventory	PnP, computer	One sample
Lukin et al (1985) [32]	Beck Depression Inventory, State-Trait Anxiety Inventory	PnP, computer	One sample crossover (2x2)
Miller et al (2002) [33]	Alcohol Use Disorder Identification Test, Alcohol Dependence Scale, Rutgers Alcohol Problem Index	PnP, online	One sample randomized
Murelle et al (1992) [34]	Michigan Alcohol Screening Test, CAGE Substance Abuse Screening Tool, Fagerstrom Tolerance Questionnaire, Center for Epidemiologic Studies Depression scale, Eating Attitudes Test, Drug Abuse Screen Test, State-Trait Anxiety Inventory	PnP, computer	One sample non random
Ogles et al (1998) [35]	Center for Epidemiologic Studies Depression scale	PnP, computer	One sample non random
Read et al (2008) [36]	PTSD Check List–Civilian Version, Traumatic Life Events Questionnaire	PnP, online	One sample non random
Schulenberg & Yutrzenka (2001) [37]	Beck Depression Inventory-II	PnP, computer	One sample crossover (4x2)
Schmitz et al (2000) [38]	Symptom Checklist 90 Revised	PnP, computer	One sample randomized

XSL•FO
RenderX

Publication	Instruments	Formats	Design
Swartz et al (2007) [39]	Center for Epidemiologic Studies Depression scale	PnP, PDA	One sample crossover (2x2)
Thorén et al (2012) [40]	Hospital Anxiety and Depression Scale	PnP, online	One sample crossover (2x2)
Thorndike et al (2011) [41]	Insomnia Severity Index	PnP, online	One sample crossover (2x2)
Vallejo et al (2007) [42]	General Health Questionnaire-28, Symptom Checklist 90 Revised	PnP, online	One sample non-random
Vallejo et al (2008) [43]	General Health Questionnaire-28, Symptom Checklist 90 Revised	PnP, online	One sample crossover (2x2)
Whitehead (2011) [44]	Hospital Anxiety and Depression Scale	PnP, online	One sample randomized
Wijndaele et al (2007) [45]	General Health Questionnaire-12, Symptom Checklist 90 Revised	PnP, online	One sample non random
Yu & Yu (2007) [46]	Center for Epidemiologic Studies Depression Scale	PnP, online	One sample randomized
Zimmerman & Martinez (2012) [47]	Clinically Useful Depression Outcome Scale	PnP, online	One sample non random
Andersson et al (2003) [48]	Hospital Anxiety and Depression Scale	PnP, online	Two samples
Hedman et al (2010) [49]	Liebowitz Social Anxiety Scale Test Self-Report, Social Phobia Scale, Social Interaction Anxiety Scale, Montgomery-Asberg Depression Rating Scale-Self-report, Beck Anxiety Inventory	PnP, online	Two samples
Le et al (2009) [50]	Panic Disorder Severity Scale, Edinburgh Postnatal Depression Scale	PnP, online	Two samples
Schmitz et al (1999) [51]	Symptom Checklist 90 Revised	PnP, computer	Two samples
Shea et al (2009) [52]	Depression Anxiety Stress Scale-21	PnP, online	Two samples

Interformat Reliability

Including the analyses conducted by the authors of this review, differences between mean scores were analyzed in 88% (29/33) of the studies; 52% (17/33) also had adequate power to detect differences of at least moderate effect size. Subscales were excluded in the analysis of interformat reliability since data from some instruments with several subscales, notably the Symptom Checklist 90 Revised (SCL-90R) and the General Health Questionnaire (GHQ-28), would have a disproportional influence on the results. Focusing on the total scores of the instruments, 62 analyses of differences between format mean scores were made, and significant differences were found in 10 (16%) of the 62. Limiting the analysis to the 17 studies with adequate power to analyze mean differences, 6 differences were found in 31 (19%) analyses. Including studies of all sample sizes, significant mean differences between scores were found in 8 of the 40 investigated instruments. The effects sizes (Cohen's *d*) of the mean score differences ranged from .14 to .98, showing that some of the effects were large.

To assess whether there was an aggregate trend in the data, both studies reporting significant differences and those reporting nonsignificant but consistent numerical differences were analyzed (ie, including analyses that showed higher but non-significant values for either condition). In total, 30 instruments or subscales reported a higher mean score for the PnP version, and 26 reported a higher mean score for the digital version. This difference in proportion was not significant (P=.26).

Correlations between format scores were reported for 28 instruments and ranged between r=.35 and r=.99. More than half of the instruments (16/33, 57%) showed strong uncontradicted correlations (>.80) between format scores, while the correlations between formats scores were ambiguous for five instruments. Strong interformat correlations have been replicated for only four instruments: Agoraphobic Cognitions Questionnaire (ACQ), Mobile Inventory (MI), Beck Depression Inventory II (BDI-II), and Montgomery-Asberg Depression Rating Scale-Self-report (MADRS-S). Two studies investigated differences in factor structure and model fit between PnP-format and digital format of questionnaires, and neither of these studies found any significant model differences. See Multimedia Appendix 4 for a review of the extracted psychometric properties.

The reported significant mean differences between format scores were not evenly distributed among the studies. Instead, many of the reported differences could be found in a small number of studies. All the 10 significant mean differences found were reported in five studies, and one particular study reported 5 (50%) of the identified differences. The five studies did not differ markedly from the other studies concerning study characteristic (eg, publication year, sample size, and quality

XSL•FO
RenderX

assessment score). The sample size in the five studies that reported inequality between administration formats ranged from 83 to 1171, with two studies having enough power to detect a small effect size. In conclusion, none of the study characteristics assessed in this study could explain the reported inequalities between formats found in these studies.

Differences in mean scores from the respective format were found in ACQ, Beck Anxiety Inventory (BAI), BDI-II (twice), Body Sensations Questionnaire (BSQ) (twice), Center for Epidemiologic Studies Depression Scale (CES-D), MI, SCL-90-R, and State-Trait Anxiety Inventory-State (STAI-S). Of these instruments, ACQ, BAI, BDI-II, CES-D, and SCL-90-R were also investigated in other studies with sufficient power to detect effects of corresponding sizes without finding significant mean differences, making results regarding interformat reliability contradictory (see Multimedia Appendix 4). BSQ and MI have either not been repeatedly investigated or been investigated in studies with insufficient power.

Test-Retest

Test-retest analysis for digital formats was conducted for 14 instruments in six studies, and the mean correlation between test occasions was *r*=.84 (SD .07, range *r*=.70-.90). The majority of the instruments (10/14, 71%) showed good test-retest reliability (>.80). Further, no significant effects of time were reported for the seven instruments analyzed with two-way ANOVA.

Internal Consistency

Internal consistency for the digital version of the instruments was reported in 24 studies and for 26 different instruments. Cronbach alpha was calculated for the digital format of an instrument in 69 instances and the mean value was .87 (SD .09, range .52-.97). A large majority of the instruments (24/26, 94%) showed adequate internal consistency (alpha>.70). Questionable internal consistencies were found for 12-Item Short Form Health Survey (SF-12 v2), Mental Component Scale, and for subscales of SCL-90-R, and the alpha results regarding Insomnia Severity Index were ambiguous.

Discussion

Principal Findings

This review aimed to investigate the interformat reliability of self-report instruments used in psychotherapy research. Studies comparing digital and PnP formats of 40 different instruments, covering various psychiatric disorders, were identified in the review process. Similar to a previous study in the somatic field, this systematic review of the literature showed that generally the reliability between digital and pen and paper versions of instruments is high [12]. The large majority of studies found adequate correlations between format scores and no significant differences between means derived from the respective format. For example, high quality studies consistently report a high interformat reliability for MADRS-S showing that this instrument can be used with confidence in online psychotherapy research. Several other well-known instruments, such as the Alcohol Use Disorder Identification Test (AUDIT), PCL-C, and Patient Health Questionnaire (PHQ-9) have also shown high interformat reliability, but the results have not yet been replicated in high quality studies. While most instruments have been investigated only once, some instruments, notably the BDI and the CES-D, have been investigated multiple times. Generally, these studies support the interformat reliability of the BDI and the CES-D. In contrast, the instruments SCL-90-R and GHQ-28 showed less than satisfactory reliability in several studies. The reasons for this are unknown, but both instruments are rather complex and contain several subscales that are designed to capture many different domains of psychological health. It is possible that either the complexity or the broader psychiatric scope of these instruments, compared to most others in this review, make them sensitive to a change of format.

Significant differences in mean scores between formats were found in 16% of all analyses. These differences were found in a small number of studies, indicating that the results may be due to study or participant characteristics rather than the properties of the instruments. The effects sizes of the mean differences ranged from small to large, indicating that the significant differences were not just a matter of high power in studies with large samples. It is noteworthy that several high powered studies did not find significant differences between format scores. At the same time, as exemplified by the largest study included in this review, by Yu & Yu [47], the effect may often be too small to be detected in smaller studies. Such small effects of instrument format would not have any major implications for most psychotherapy research but may, for example, affect the results of prevalence studies. Also, some of the differences found between format means were of medium or large effect sizes, which implies that study design may affect the results substantially. If researchers are not careful and meticulous in designing their studies, the results from computer-based psychotherapy may potentially not be comparable to that of traditional psychotherapy.

The range of correlations between formats was wide, from .35 to .99. The lowest correlations were found in certain subscales of the general health questionnaires, scales that overall show lower reliability, and the STAI-S, an instrument that explicitly measures the current condition and may thus be very sensitive to time effects. Interestingly, significant format differences were reported for the BDI, MI, and ACQ, while these instruments at the same time reported high correlations. This underscores the importance of not solely relying on correlations when assessing interformat reliability.

When performing a systematic review, there is always the risk of selection bias when searching for and including relevant studies. This includes publication bias, which automatically narrows the range of studies that can be included but is also relevant when perusing reference lists of identified papers for additional studies. However, these risks may be somewhat lower in the case of studies on interformat reliability since both positive and negative results should draw attention from publishers. Still, the risk of bias in the included studies in any review should not be underestimated.

The majority of studies were of high scientific quality, using adequate designs and statistical analyses. However, only half of the studies had an adequate sample size to detect mean

XSL•FO
RenderX

differences with a medium effect size and power calculations were seldom reported. Also, the time interval between data collections, arguably a very important factor, was not reported in all studies using a crossover design. Possible interval differences may thus explain some of the variance found between studies. Future studies should focus on improving the methodological quality by increasing the sample size to achieve sufficient power. Another recommendation is that studies report both interformat correlation and differences between format means.

Knowledge about factors that may affect interformat reliability is still limited. One potential factor is the characteristics of the digital format itself. The layout, user interface, etc, are likely to affect the score of an instrument in some way, at least if these characteristics are markedly different from the PnP version. It is valuable to know the degree of similarity between the PnP and the digital version in order to assess this potential effect. Regrettably, few studies report what adaptations are made when transferring the instrument to the new administration format.

In this review, only three studies used digital formats other than computer/Internet: Cooke et al [22] used a palm device, Swartz et al [39] used a personal digital assistant, and Bush et al [17] used online and smartphone formats. None of these studies reported any significant differences of mean values between formats, at least indicating that interformat reliability could be stable over different forms of digital platforms.

Further, respondents' reactions to the digital format may affect results if the digital medium is perceived as different, for example, as more secure or anonymous than the traditional pen and paper medium. In research on survey methodology, a number of studies have shown that data collected on the Internet can be equivalent to data collected with traditional methods [53-55]. While there are no empirical studies investigating how participants react to instruments in psychotherapy research, the conclusions in the present review are in line with these results.

Limitations

This study has some limitations. Only instruments typically used in psychotherapy research were included. Psychiatric symptom scales of the questionnaire type may be less sensitive to administration differences than some other measurements, such as VAS-scales or behavioral diaries. It is thus unclear whether the results of this study could be generalized to these other types of measurements as well. Further, no effort was made to contact authors for additional data if relevant information was missing in the articles included. Since several studies failed to report variables that were investigated in this study, contact with authors may have contributed more data. Finally, since only reference lists from included studies were perused, it is possible that we missed studies that were cited in papers that we reviewed but excluded. A strength of the current study was the effort to find and include older studies as well as studies with different designs.

While this review focused on reliability of digital instruments, future studies could also investigate aspects of validity. It could be argued that instruments showing adequate interformat reliability do not need to confirm the validity for digital use if validity is already established for their pen and paper versions. This is, however, an empirical question and could be important to consider. To our knowledge, few studies have investigated the validity of digital instruments.

In general, instruments used in Internet-based psychotherapy research show high interformat reliability and can be used with confidence. There are also some signs that the factor structure is not affected by delivery format. There is, however, still a need for well-designed and high powered studies investigating the most widely used instruments, such as the PHQ-9. While the use of mobile technology will increase, very few studies have investigated instruments administrated through a smartphone, tablet computer, or similar device. Future studies could thus focus on these platforms.

However, even within platforms, such as personal computer or smartphone, there is almost limitless variation when it comes to instrument presentation. The format per se may be much less important than the specific presentation. Different presentations and adaptations of instruments could, in the future, be investigated experimentally to identify factors that influence interformat reliability. Since few studies report in any detail what adaptations they have made of the digital instruments, effect of presentation is still largely unknown. One of the benefits of digital instruments is the possibility to design smart questionnaires that adapt to the respondent's answers. While outside the scope of the present review, such development may be more relevant for clinical care in the future [56].

Conclusion

This review concludes that, while instruments in most studies show high interformat reliability, there are some exceptions, and it is still unclear if these exceptions are due to psychometric properties of specific instrument or to study properties. In general, instruments used in psychotherapy research seem to be robust over administration formats. Future studies should increase sample sizes and both investigate and clearly report how digital adaptation of instruments are made.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Study quality assessment.

[PDF File (Adobe PDF File), 98KB - jmir_v16i12e268_app1.pdf]

XSL•FO
RenderX

Multimedia Appendix 2

PRISMA flow chart.

[PDF File (Adobe PDF File), 57KB - jmir_v16i12e268_app2.pdf]

Multimedia Appendix 3

Study and participant characteristics.

[PDF File (Adobe PDF File), 42KB - jmir_v16i12e268_app3.pdf]

Multimedia Appendix 4

Psychometric results from all studies.

[PDF File (Adobe PDF File), 149KB - jmir_v16i12e268_app4.pdf]

References

1. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. Cogn Behav Ther 2009;38(4):196-205. [doi: 10.1080/16506070903318960] [Medline: 20183695]
2. Cuijpers P, Marks IM, van Straten A, Cavanagh K, Gega L, Andersson G. Computer-aided psychotherapy for anxiety disorders: a meta-analytic review. Cogn Behav Ther 2009;38(2):66-82. [doi: 10.1080/16506070802694776] [Medline: 20183688]
3. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
4. Coons SJ, Gwaltney CJ, Hays RD, Lundy JJ, Sloan JA, Revicki DA, ISPOR ePRO Task Force. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO Good Research Practices Task Force report. Value Health 2009 Jun;12(4):419-429. [doi: 10.1111/j.1524-4733.2008.00470.x] [Medline: 19900250]
5. Thorndike FP, Carlbring P, Smyth FL, Magee JC, Gonder-Frederick L, Ost L, et al. Web-based measurement: Effect of completing single or multiple items per webpage. Computers in Human Behavior 2009 Mar;25(2):393-401. [doi: 10.1016/j.chb.2008.05.006]
6. Tourangeau R. Spacing, Position, and Order: Interpretive Heuristics for Visual Features of Survey Questions. Public Opinion Quarterly 2004 Sep 01;68(3):368-393. [doi: 10.1093/poq/nfh035]
7. Hunter J, Corcoran K, Leeder S, Phelps K. Is it time to abandon paper? The use of emails and the Internet for health services research--a cost-effectiveness and qualitative study. J Eval Clin Pract 2013 Oct;19(5):855-861. [doi: 10.1111/j.1365-2753.2012.01864.x] [Medline: 22640009]
8. Schulenberg SE, Yutrzenka BA. The equivalence of computerized and paper-and-pencil psychological instruments: implications for measures of negative affect. Behav Res Methods Instrum Comput 1999 May;31(2):315-321. [Medline: 10495816]
9. Joinson A. Social desirability, anonymity, and Internet-based questionnaires. Behav Res Methods Instrum Comput 1999 Aug;31(3):433-438. [Medline: 10502866]
10. Ferriter M. Computer aided interviewing and the psychiatric social history. Social Work and Social Sciences Review 1993;4(3):255-263.
11. Rogers SM, Willis G, Al-Tayyib A, Villarroel MA, Turner CF, Ganapathi L, et al. Audio computer assisted interviewing to measure HIV risk behaviours in a clinic population. Sex Transm Infect 2005 Dec;81(6):501-507 [FREE Full text] [doi: 10.1136/sti.2004.014266] [Medline: 16326855]
12. Gwaltney CJ, Shields AL, Shiffman S. Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: a meta-analytic review. Value Health 2008;11(2):322-333. [doi: 10.1111/j.1524-4733.2007.00231.x] [Medline: 18380645]
13. Coaley K. An introduction to psychological assessment and psychometrics. London: SAGE Publications Ltd; 2014.
14. Streiner D, Norman G. Health measurement scales: a practical guide to their development and use. Oxford: Oxford university press; 2008.
15. Austin DW, Carlbring P, Richards JC, Andersson G. Internet Administration of Three Commonly Used Questionnaires in Panic Research: Equivalence to Paper Administration in Australian and Swedish Samples of People With Panic Disorder. International Journal of Testing 2006 Mar;6(1):25-39. [doi: 10.1207/s15327574ijt0601_2]
16. Brock RL, Barry RA, Lawrence E, Dey J, Rolffs J. Internet administration of paper-and-pencil questionnaires used in couple research: assessing psychometric equivalence. Assessment 2012 Jun;19(2):226-242. [doi: 10.1177/1073191110382850] [Medline: 20881103]

XSL•FO
RenderX

17. Bush NE, Skopp N, Smolenski D, Crumpton R, Fairall J. Behavioral screening measures delivered with a smartphone app: psychometric properties and user preference. J Nerv Ment Dis 2013 Nov;201(11):991-995. [doi: 10.1097/NMD.0000000000000039] [Medline: 24177488]
18. Butler N, Newton T, Slade P. Validation of a computerized version of the scans questionnaire. Int. J. Eat. Disord 1989 Mar;8(2):239-241. [doi: 10.1002/1098-108X(198903)8:2<239::AID-EAT2260080215>3.0.CO;2-1]
19. Carlbring P, Brunt S, Bohman S, Austin D, Richards J, Öst L, et al. Internet vs. paper and pencil administration of questionnaires commonly used in panic/agoraphobia research. Computers in Human Behavior 2007 May;23(3):1421-1434. [doi: 10.1016/j.chb.2005.05.002]
20. Chan-Pensley E. Alcohol-Use Disorders Identification Test: a comparison between paper and pencil and computerized versions. Alcohol Alcohol 1999;34(6):882-885 [FREE Full text] [Medline: 10659724]
21. Coles ME, Cook LM, Blake TR. Assessing obsessive compulsive symptoms and cognitions on the internet: evidence for the comparability of paper and Internet administration. Behav Res Ther 2007 Sep;45(9):2232-2240. [doi: 10.1016/j.brat.2006.12.009] [Medline: 17306222]
22. Cook IA, Balasubramani GK, Eng H, Friedman E, Young EA, Martin J, et al. Electronic source materials in clinical research: acceptability and validity of symptom self-rating in major depressive disorder. J Psychiatr Res 2007 Nov;41(9):737-743. [doi: 10.1016/j.jpsychires.2006.07.015] [Medline: 17275840]
23. Fortson BL, Scotti JR, Del Ben KS, Chen YC. Reliability and validity of an Internet traumatic stress survey with a college student sample. J Trauma Stress 2006 Oct;19(5):709-720. [doi: 10.1002/jts.20165] [Medline: 17075908]
24. George CE, Lankford J, Wilson SE. The effects of computerized versus paper-and-pencil administration on measures of negative affect. Computers in Human Behavior 1992 Jan;8(2-3):203-209. [doi: 10.1016/0747-5632(92)90004-X]
25. Glaze R, Cox JL. Validation of a computerised version of the 10-item (self-rating) Edinburgh Postnatal Depression Scale. J Affect Disord 1991;22(1-2):73-77. [Medline: 1880310]
26. Herrero J, Meneses J. Short Web-based versions of the perceived stress (PSS) and Center for Epidemiological Studies-Depression (CESD) Scales: a comparison to pencil and paper responses among Internet users. Computers in Human Behavior 2006 Sep;22(5):830-846. [doi: 10.1016/j.chb.2004.03.007]
27. Hirai M, Vernon LL, Clum GA, Skidmore ST. Psychometric Properties and Administration Measurement Invariance of Social Phobia Symptom Measures: Paper-Pencil vs. Internet Administrations. J Psychopathol Behav Assess 2011 Sep 8;33(4):470-479. [doi: 10.1007/s10862-011-9257-2]
28. Holländare F, Gahnström A, Nieminen A, Engström I. Can BDI-II and MADRS-S be Transferred to Online Use Without Affecting its Psychometric Properties? EJAP 2009 Jan 30;4(2):-. [doi: 10.7790/ejap.v4i2.122]
29. Holländare F, Andersson G, Engström I. A comparison of psychometric properties between internet and paper versions of two depression instruments (BDI-II and MADRS-S) administered to clinic patients. J Med Internet Res 2010;12(5):e49 [FREE Full text] [doi: 10.2196/jmir.1392] [Medline: 21169165]
30. Kurt R, Bogner HR, Straton JB, Tien AY, Gallo JJ. Computer-assisted assessment of depression and function in older primary care patients. Comput Methods Programs Biomed 2004 Feb;73(2):165-171 [FREE Full text] [Medline: 14757259]
31. Lankford J, Bell RW, Elias JW. Computerized versus standard personality measures: Equivalency, computer anxiety, and gender differences. Computers in Human Behavior 1994 Dec;10(4):497-510. [doi: 10.1016/0747-5632(94)90043-4]
32. Lukin ME, Dowd ET, Plake BS, Kraft RG. Comparing computerized versus traditional psychological assessment. Computers in Human Behavior 1985 Jan;1(1):49-58. [doi: 10.1016/0747-5632(85)90006-8]
33. Miller ET, Neal DJ, Roberts LJ, Baer JS, Cressler SO, Metrik J, et al. Test-retest reliability of alcohol measures: is there a difference between internet-based assessment and traditional methods? Psychol Addict Behav 2002 Mar;16(1):56-63. [Medline: 11934087]
34. Murrelle L, Bulger JD, Ainsworth BE, Holliman SC, Bulger DW. Computerized mental health risk appraisal for college students: user acceptability and correlation with standard pencil-and-paper questionnaires. Am J Health Promot 1992;7(2):90-92. [Medline: 10146792]
35. Ogles BM, France CR, Lunnen KM, Bell MT, Goldfarb M. Computerized depression screening and awareness. Community Ment Health J 1998 Feb;34(1):27-38. [Medline: 9559238]
36. Read JP, Farrow SM, Jaanimägi U, Ouimette P. Assessing trauma and traumatic stress via the Internet: Measurement equivalence and participant reactions. Traumatology 2009;15(1):94-102. [doi: 10.1177/1534765608325121]
37. Schulenberg SE, Yutrzenka BA. Equivalence of computerized and conventional versions of the Beck Depression Inventory-II (BDI-II). Curr Psychol 2001 Sep;20(3):216-230. [doi: 10.1007/s12144-001-1008-1]
38. Swartz RJ, de Moor C, Cook KF, Fouladi RT, Basen-Engquist K, Eng C, et al. Mode effects in the center for epidemiologic studies depression (CES-D) scale: personal digital assistant vs. paper and pencil administration. Qual Life Res 2007 Jun;16(5):803-813. [doi: 10.1007/s11136-006-9158-0] [Medline: 17295102]
39. Thorén ES, Andersson G, Lunner T. The use of research questionnaires with hearing impaired adults: online vs. paper-and-pencil administration. BMC Ear Nose Throat Disord 2012;12:12 [FREE Full text] [doi: 10.1186/1472-6815-12-12] [Medline: 23107440]

XSL•FO
RenderX

40. Thorndike FP, Ritterband LM, Saylor DK, Magee JC, Gonder-Frederick LA, Morin CM. Validation of the insomnia severity index as a web-based measure. Behav Sleep Med 2011;9(4):216-223. [doi: 10.1080/15402002.2011.606766] [Medline: 22003975]
41. Vallejo MA, Jordán CM, Díaz M, Comeche MI, Ortega J. Psychological assessment via the internet: a reliability and validity study of online (vs paper-and-pencil) versions of the General Health Questionnaire-28 (GHQ-28) and the Symptoms Check-List-90-Revised (SCL-90-R). J Med Internet Res 2007;9(1):e2 [FREE Full text] [doi: 10.2196/jmir.9.1.e2] [Medline: 17478411]
42. Vallejo MA, Mañanes G, Isabel Comeche MA, Díaz M. Comparison between administration via Internet and paper-and-pencil administration of two clinical instruments: SCL-90-R and GHQ-28. J Behav Ther Exp Psychiatry 2008 Sep;39(3):201-208. [doi: 10.1016/j.jbtep.2007.04.001] [Medline: 17573039]
43. Whitehead L. Methodological issues in Internet-mediated research: a randomized comparison of internet versus mailed questionnaires. J Med Internet Res 2011;13(4):e109 [FREE Full text] [doi: 10.2196/jmir.1593] [Medline: 22155721]
44. Wijndaele K, Matton L, Duvigneaud N, Lefevre J, Duquet W, Thomis M, et al. Reliability, equivalence and respondent preference of computerized versus paper-and-pencil mental health questionnaires. Computers in Human Behavior 2007 Jul;23(4):1958-1970. [doi: 10.1016/j.chb.2006.02.005]
45. Yu SC, Yu MN. Comparison of Internet-based and paper-based questionnaires in Taiwan using multisample invariance approach. Cyberpsychol Behav 2007 Aug;10(4):501-507. [doi: 10.1089/cpb.2007.9998] [Medline: 17711357]
46. Zimmerman M, Martinez JH. Web-based assessment of depression in patients treated in clinical practice: reliability, validity, and patient acceptance. J Clin Psychiatry 2012 Mar;73(3):333-338. [doi: 10.4088/JCP.10m06519] [Medline: 22154899]
47. Andersson G, Kaldo-Sandström V, Ström L, Strömgren T. Internet administration of the Hospital Anxiety and Depression Scale in a sample of tinnitus patients. J Psychosom Res 2003 Sep;55(3):259-262. [Medline: 12932800]
48. Hedman E, Ljótsson B, Rück C, Furmark T, Carlbring P, Lindefors N, et al. Internet administration of self-report measures commonly used in research on social anxiety disorder: A psychometric evaluation. Computers in Human Behavior 2010 Jul;26(4):736-740. [doi: 10.1016/j.chb.2010.01.010]
49. Le HN, Perry DF, Sheng X. Using the internet to screen for postpartum depression. Matern Child Health J 2009 Mar;13(2):213-221. [doi: 10.1007/s10995-008-0322-8] [Medline: 18278545]
50. Schmitz N, Hartkamp N, Brinschwitz C, Michalek S. Computerized administration of the Symptom Checklist (SCL-90-R) and the Inventory of Interpersonal Problems (IIP-C) in psychosomatic outpatients. Psychiatry Res 1999 Oct 11;87(2-3):217-221. [Medline: 10579554]
51. Schmitz N, Hartkamp N, Brinschwitz C, Michalek S, Tress W. Comparison of the standard and the computerized versions of the Symptom Check List (SCL-90-R): a randomized trial. Acta Psychiatr Scand 2000 Aug;102(2):147-152. [Medline: 10937788]
52. Shea TL, Tennant A, Pallant JF. Rasch model analysis of the Depression, Anxiety and Stress Scales (DASS). BMC Psychiatry 2009;9:21 [FREE Full text] [doi: 10.1186/1471-244X-9-21] [Medline: 19426512]
53. Gosling SD, Vazire S, Srivastava S, John OP. Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. Am Psychol 2004;59(2):93-104. [doi: 10.1037/0003-066X.59.2.93] [Medline: 14992636]
54. Leece P, Bhandari M, Sprague S, Swiontkowski MF, Schemitsch EH, Tornetta P, et al. Internet versus mailed questionnaires: a controlled comparison. J Med Internet Res 2004 Oct 29;6(4):e39 [FREE Full text] [doi: 10.2196/jmir.6.4.e39] [Medline: 15631963]
55. Tourangeau R, Yan T. Sensitive questions in surveys. Psychol Bull 2007;133(5):859-883. [doi: 10.1037/0033-2909.133.5.859]
56. Ruland C, Røslien J, Bakken S, Kristiansen J. Comparing tailored computerized symptom assessments to interviews and questionnaires. AMIA Annu Symp Proc 2006:1081 [FREE Full text] [Medline: 17238700]

Abbreviations

ACQ: Agoraphobic Cognitions Questionnaire
ANOVA: analysis of variance
AUDIT: Alcohol Use Disorder Identification Test
BAI: Beck Anxiety Inventory
BDI: Beck Depression Inventory
BSQ: Body Sensations Questionnaire
CES-D: Center for Epidemiologic Studies Depression scale
GHQ: General Health Questionnaire
MADRS-S: Montgomery–Asberg Depression Rating Scale Self-report
MI: Mobility Inventory
PCL-C: PTSD Check List–Civilian Version
PHQ: Patient Health Questionnaire
PnP: pen and paper

XSL•FO
RenderX

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SCL-90-R: Symptom Checklist 90 Revised
SF12V2: Short Form (12) Health Survey Version Two
STAI-S: State-Trait Anxiety Inventory–State
STAI-T: State-Trait Anxiety Inventory–Trait
VAS: Visual Analogue Scale

Edited by G Eysenbach; submitted 13.03.14; peer-reviewed by MA Johansen, C Lalanne, MA Vallejo Pareja; comments to author 16.07.14; revised version received 12.08.14; accepted 16.08.14; published 03.12.14.

Please cite as:
Alfonsson S, Maathz P, Hursti T
Interformat Reliability of Digital Psychiatric Self-Report Questionnaires: A Systematic Review
J Med Internet Res 2014;16(12):e268
URL: http://www.jmir.org/2014/12/e268/
doi:10.2196/jmir.3395
PMID:25472463

©Sven Alfonsson, Pernilla Maathz, Timo Hursti. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 03.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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

Nonprobability Web Surveys to Measure Sexual Behaviors and Attitudes in the General Population: A Comparison With a Probability Sample Interview Survey

Bob Erens[1,2], MA; Sarah Burkill[2], MSc; Mick P Couper[3], PhD; Frederick Conrad[3], PhD; Soazig Clifton[2,4], BSc (Hons); Clare Tanton[2], PhD; Andrew Phelps[4], BA; Jessica Datta[1], MSc; Catherine H Mercer[2], PhD; Pam Sonnenberg[2], PhD; Philip Prah[2], MSc; Kirstin R Mitchell[1], PhD; Kaye Wellings[1], FRCOG; Anne M Johnson[2], MD; Andrew J Copas[2], PhD

[1]London School of Hygiene & Tropical Medicine, London, United Kingdom
[2]University College London, London, United Kingdom
[3]University of Michigan, Ann Arbor, MI, United States
[4]NatCen Social Research, London, United Kingdom

Corresponding Author:
Bob Erens, MA
London School of Hygiene & Tropical Medicine
15-17 Tavistock Place
London, WC1H 9SH
United Kingdom
Phone: 44 (0)20 7927 2784
Fax: 44 (0)20 7927 2701
Email: bob.erens@lshtm.ac.uk

Abstract

Background: Nonprobability Web surveys using volunteer panels can provide a relatively cheap and quick alternative to traditional health and epidemiological surveys. However, concerns have been raised about their representativeness.

Objective: The aim was to compare results from different Web panels with a population-based probability sample survey (n=8969 aged 18-44 years) that used computer-assisted self-interview (CASI) for sensitive behaviors, the third British National Survey of Sexual Attitudes and Lifestyles (Natsal-3).

Methods: Natsal-3 questions were included on 4 nonprobability Web panel surveys (n=2000 to 2099), 2 using basic quotas based on age and sex, and 2 using modified quotas based on additional variables related to key estimates. Results for sociodemographic characteristics were compared with external benchmarks and for sexual behaviors and opinions with Natsal-3. Odds ratios (ORs) were used to express differences between the benchmark data and each survey for each variable of interest. A summary measure of survey performance was the average absolute OR across variables. Another summary measure was the number of key estimates for which the survey differed significantly (at the 5% level) from the benchmarks.

Results: For sociodemographic variables, the Web surveys were less representative of the general population than Natsal-3. For example, for men, the average absolute OR for Natsal-3 was 1.14, whereas for the Web surveys the average absolute ORs ranged from 1.86 to 2.30. For all Web surveys, approximately two-thirds of the key estimates of sexual behaviors were different from Natsal-3 and the average absolute ORs ranged from 1.32 to 1.98. Differences were appreciable even for questions asked by CASI in Natsal-3. No single Web survey performed consistently better than any other did. Modified quotas slightly improved results for men, but not for women.

Conclusions: Consistent with studies from other countries on less sensitive topics, volunteer Web panels provided appreciably biased estimates. The differences seen with Natsal-3 CASI questions, where mode effects may be similar, suggest a selection bias in the Web surveys. The use of more complex quotas may lead to some improvement, but many estimates are still likely to differ. Volunteer Web panels are not recommended if accurate prevalence estimates for the general population are a key objective.

(J Med Internet Res 2014;16(12):e276) doi:10.2196/jmir.3382

XSL•FO
RenderX

KEYWORDS

Internet survey; Web survey; survey methods; sampling bias; selection bias; sexual behavior

Introduction

Over the past decade, there has been dramatic growth in both Europe and the United States in the use of Web surveys for market research and opinion polling. However, the Web has not been widely used for collecting epidemiological or health surveillance data (or for academic research more generally) despite increasing interest in doing so [1-12]. This is largely due to the reliance on volunteer Web panels (at least when examining the general population) and well-founded concerns about the representativeness of such nonprobability Web surveys that rely on these panels [13].

Volunteer Web panels typically include hundreds of thousands of potential participants who have signed up to participate in Web surveys, often for a small incentive (eg, a payment or points that can be redeemed for goods). Although there are a few Web panels in the United States (the GfK Knowledge Panel) and in Europe (eg, the Longitudinal Internet Study for the Social sciences [LISS] panel in the Netherlands and the German Internet Panel) selected by using probability-based sampling methods, most Web panels are made up of self-selected volunteers who are recruited using a variety of methods (eg, through email databases, visitors to websites, online advertisements). Aside from sampling issues, the main concerns about Web panels include coverage bias because individuals or households without access to the Internet are excluded, and nonresponse bias because response rates to Web surveys are often very low if reported at all [13-16].

Despite these concerns, the use of Web surveys is likely to continue to increase in Britain and elsewhere because they purportedly allow efficient, relatively cheap, and quick data collection [14,17-20], advantages which will increasingly appeal to academic and government researchers faced with rising costs for traditional face-to-face and telephone interview surveys at a time of constrained budgets (eg, see the UK government's rationale for moving the Community Life Survey to Web data collection [21]) . Web surveys may even be advantageous for certain types of studies, such as those concerned with sensitive behaviors, because interviewers are not present and the greater privacy provided by Web surveys may lead to higher reporting of socially undesirable behaviors [3,8,10,12,19,22,23]. Therefore, it is important to continue to evaluate the representativeness of volunteer Web panels and the circumstances in which they can be used, and to attempt to develop approaches and estimation methods to improve the robustness of Web panel data.

This paper makes a number of contributions to the evidence base by comparing results from 4 volunteer Web panel surveys with those from the third British National Survey of Sexual Attitudes and Lifestyles (Natsal-3), a probability sample interview survey. Firstly, unlike much of the previous research on this topic, which generally compares results for voting intentions or opinion questions, Natsal-3 includes primarily behavioral measures along with a few opinion questions. Secondly, Natsal-3 includes a lengthy computer-assisted self-interview (CASI) component that, like a Web survey, requires participants to enter their answers directly onto a computer. Hence, this study makes a unique contribution by comparing results for sensitive behaviors between Web surveys and a CASI probability sample survey. Thirdly, the study included an experiment that modified the quota controls for 2 of the Web surveys to see whether this would improve the representativeness of their results. Fourthly, although there has been considerable research in the United States and Europe comparing results from Web panel surveys with those from traditional face-to-face, telephone, and mail surveys (much of which is summarized in the 2010 report from the American Association for Public Opinion Research [AAPOR] [13]), there has been very little published research on this topic for British Web panels, despite all the major British market research agencies maintaining large databases of Web panel members.

Methods

Data Collection

The 4 Web surveys were carried out by 3 well-known survey organizations in the United Kingdom, each involved in social and market research, and each with large volunteer Web panels. Results from the Web surveys were compared with Natsal-3 results and, for a limited number of variables, with external data. Our focus was not on the individual panels and confidentiality agreements with the companies prohibit us from identifying them.

Natsal is one of the largest surveys on sexual behavior in the world, having interviewed 18,876 adults in 1990-1991 (Natsal-1), 12,110 adults in 1999-2001 (Natsal-2), and 15,162 adults in 2010-2012 (Natsal-3). Natsal-3 involved a stratified, clustered probability sample design and an interview with 1 randomly selected adult aged 16-74 years by a trained interviewer in the participant's own home. Median interview length was 53 minutes, split (approximately equally) between a face-to-face computer-assisted personal interview (CAPI) and CASI for the most sensitive questions. Multimedia Appendix 1 shows which questionnaire sections in Natsal-3 were asked in CAPI and which were asked in CASI. Fieldwork was carried out between September 2010 and August 2012, and achieved a response rate of 57.7% (similar to Natsal-2 for the comparable age group) [24].

A subset of approximately 130 Natsal-3 questions was included on the Web surveys, which took approximately 20 minutes on average to complete. The Web survey questions were exactly the same as those asked in Natsal-3, except where changes in format were required (eg, where show cards were used in Natsal-3). A version of the Web survey questionnaire is included in Multimedia Appendix 2. The age range included in the Web surveys was restricted to adults aged 18-44 years (n=8969) and all analyses in this paper are restricted to that age range.

XSL•FO
RenderX

Two of the Web surveys set "basic" quotas on variables which are central to measuring sexual behavior (age and partnership status), whereas the other 2 involved "modified" quotas with additional variables not typically used in setting quota controls (eg, opinion variables). Given that the purpose of the study was to see if the modified quotas would bring the Web survey estimates closer to Natsal-3, the modified quotas were set using distributions from Natsal-3 for the relevant variables.

Web survey basic quota 1 (WS-B1, carried out by Company A) and Web survey basic quota 2 (WS-B2, carried out by Company B) aimed to achieve samples of approximately 2000 cases with basic quotas set for age group (18-24, 25-34, and 35-44 years) within sex, partnership status (married/living as married vs all others) within sex, and region (London vs rest of Britain). The quotas for age within sex and region were set by reference to Office for National Statistics (ONS) midyear 2010 population estimates, whereas the quota for partnership status within sex used data from the 2009 British Labour Force Survey [25]. The quotas are shown in Table 1.

Table 1. Description of the surveys included in the study.

Sample type and survey	Quotas	Achieved sample size, N (18-44 years)	Variables used for weighting[a]
Basic quota			
WS-B1 (Company A)	Age within sex	2099	Age within sex
	Partnership status within sex		
	Region		
WS-B2 (Company B)	Age within sex	2000	Age within sex
	Partnership status within sex		
	Region		
Modified quota			
WS-M1 (Company B)	Age within sex	2000	Age within sex
	Partnership status within sex		
	Region		
	Age left full-time education		
	Any under 18s in household		
WS-M2 (Company C)	Age within sex	2021	Age within sex
	Partnership status within sex		
	Frequency of drinking alcohol		
	Age left full-time education		
	Attitude to same-gender sex		
	Household size		
Probability			
Natsal-3	Not applicable	8969	Age within sex
			Region

[a] From 2011 Census.

Web survey modified quota 1 (WS-M1), also carried out by Company B, used modified quota controls. These were determined by identifying which sociodemographic characteristics available in Natsal-3 and that Company B also had available for their Web panel members were significantly associated with key behavioral measures (including those in Multimedia Appendix 3). Then, we examined how WS-B2 and Natsal-3 differed in terms of these identified characteristics. For this second process, the WS-B2 dataset was combined with the Natsal-3 dataset (because Natsal-3 fieldwork was not yet complete, this analysis used only the first full year of Natsal-3 fieldwork and included 4459 participants aged 18-44 years). An indicator of whether the participant belonged to Natsal-3 or WS-B2 was then used as the dependent variable in a forward stepwise logistic regression model selection process using $P<.05$ in STATA 12.0 (StataCorp LP, College Station, TX, USA). The variables included in the process were age finished full-time education, participant's current economic activity, whether there were any residents younger than 18 years living in the participant's household, household size, and the area-based Index of Multiple Deprivation (IMD) score for the participant's postcode [26]. The first 3 of these listed variables were selected, but because the first 2 are strongly correlated, age finished full-time education was chosen (as the more statistically significant) along with any residents younger than 18 years in the household as factors to modify the basic quotas for WS-M1;

XSL•FO
RenderX

quotas were set so that the distribution for these 2 variables would match those obtained in Natsal-3.

Because market research agencies tend to collect very limited information on all their Web panel members, for Web survey modified quota 2 (WS-M2; carried out by Company C) the aim was to identify a number of variables associated with key measures that are *not* normally available for panel members, but which could be obtained for a large subsample of the panel by including questions on an initial Web omnibus (ie, regular multipurpose) survey. Using Natsal-3 data, significant associations in bivariate analysis between key behaviors (including those in Multimedia Appendix 3) and sociodemographic and attitudinal variables were examined to generate a shortlist of possible questions to be included on Company C's regular Web omnibus survey. The sociodemographic and attitudinal variables with the highest number of significant associations with key behaviors and which remained significant after adjusting for age, partnership status, and region in the logistic regression analysis were identified as potential variables to be used as additional quotas. WS-M2 then proceeded in 2 stages. First, 6 additional questions (on current smoking status, frequency of drinking alcohol, age completed full-time education, tenure, and attitude toward same-gender sex and abortion) were included on Company C's Web omnibus. Although the original target was to collect data for these 6 variables for approximately 30,000 members of Company C's Web panel, data were collected for only 9176 panel members within the 18-44 age range. The second stage used a forward stepwise model with Natsal-3 or the omnibus survey as a binary outcome to indicate which of the 6 variables, along with some basic sociodemographic variables held by the agency, best modeled the difference between the 2 surveys. Four variables came out as highly significant (frequency of drinking alcohol, age completed education, attitude toward same-gender sex, and household size) and were selected as additional factors to form quotas (along with age and partnership status within sex) for WS-M2. Quotas were set so that the distributions of these 4 variables would match those obtained in Natsal-3. Because it was not possible to fill all the quotas using only the 9176 panel members for whom this additional information was now available, more panel members were invited to participate in WS-M2. Those who had not completed the Web omnibus were filtered into the relevant quotas by answering the additional questions before starting the main questionnaire (and were excluded from the main questionnaire if their quota was already filled).

By the time the full Natsal-3 dataset became available for analysis (in early 2013), data from the 2011 UK population census were also available. Natsal-3 and all Web surveys were poststratified to 2011 census figures for age within sex (which differed slightly from the 2010 midyear population estimates). Natsal-3 data were also weighted by region, but we did not weight the Web surveys by region because the agencies did not collect regional data on the same basis. In the event, weighting by region did not greatly affect the estimates.

Participation in Natsal-3 and the Web surveys was based on fully informed consent. The study was approved by the Oxfordshire Research Ethics Committee A (reference number 12/SC/0070).

Analysis Methods

Bias regarding participant sociodemographic characteristics was assessed by comparing the estimates from the 4 Web surveys and Natsal-3 with external benchmark data from the 2011 UK population census; the ONS Integrated Household Survey (IHS) for 2011 (except for sexual identity which comes from IHS April 2011-March 2012), a large annual survey of approximately 400,000 individuals used to produce official statistics [27]; or from the 2010 National Travel Survey for benchmarks on holding a driving license [28]. Comparisons were made for a number of key behavioral measures reported in the initial series of Natsal-3 papers published in *The Lancet* [29-32] and also key attitudinal measures. For these measures, Natsal-3 was treated as the benchmark because its results have been widely used within government and academia, and because it is the only probability-based survey measuring these topics in the British population. No independent benchmarks exist for the behavioral and attitudinal variables.

The difference between the benchmark data and each survey for each outcome variable of interest is expressed as the odds ratio (OR). These ORs were obtained from binary or ordinal logistic regression depending on the nature of the outcome (no outcome was continuous). Nonordered categorical variables were reduced to binary form to avoid multiple ORs for 1 variable. The ORs are presented with 95% confidence intervals (CIs). One summary measure of the performance of a survey was the number of key estimates for which the survey differed significantly from the benchmark at the 5% level. Another measure of overall bias for a survey was the average absolute OR across outcomes, where the absolute OR for an OR less than 1 was calculated as 1/OR (eg, an OR of 0.5 is treated as 2.0). Average absolute ORs were also calculated separately for CAPI and CASI and for behavior and opinion questions because the performance of the Web surveys could differ by these question types. Absolute OR was selected in preference to absolute difference because we felt it better reflected the importance of differences across the range of prevalence from rare to common.

Generalized estimating equations were used to test whether the 2 basic quota Web surveys were consistent (ie, gave the same responses to each question). Robust standard errors to reflect the within-person clustering of outcome responses were used. To assess whether either approach to the use of modified quotas resulted in more accurate estimates, the average absolute OR across the outcomes for the 2 basic Web surveys (WS-B1 and WS-B2) was compared with the corresponding average OR for each modified Web survey (WS-M1 and WS-M2). The bootstrap method [33], resampling participants in each Web survey 100 times, was used to obtain a standard error for the average absolute OR for each modified survey and for the 2 basic surveys combined. This focused on the behavior and opinion outcomes because these were of central interest. The uncertainty in Natsal-3 estimates was not considered because it was not relevant to the comparison of Web surveys. The average absolute OR was bounded by 1 and unlikely to be normally

XSL•FO
RenderX

distributed, so CIs are not provided but approximate tests for difference between average absolute ORs were performed under the assumption of normality.

Results

Meeting the Web Survey Quotas

The age-sex quotas were generally met by all 4 Web surveys, except for young men aged 18-24 years which was undertarget in 2 of the surveys; in WS-B1, 211 of the target number of 267 (79.0%) was achieved and in WS-M2 it was 228 out of 267 (85.4%).

Even with Web panels containing hundreds of thousands of members, both Companies B and C had difficulties meeting the modified quotas. Both WS-M1 and WS-M2 could not find enough people who finished their education before age 17 years. WS-M2 also fell short of the quotas for large (≥4 person) households, infrequent drinkers of alcohol (those who drink less than once a week), and 1 attitude question (tolerance of same-sex relationships). These quotas had to be relaxed to achieve the target of 2000 completed questionnaires (Multimedia Appendix 4).

Comparing Participant Characteristics With External Benchmarks

Estimates regarding participant characteristics from all 4 Web surveys and Natsal-3 were compared with external benchmarks to assess bias. Variables used as quotas for any of the Web surveys were not included, leaving 6 variables to be compared: housing tenure, current economic activity, ethnicity, self-assessed general health, whether the participant had a driving licence valid in the United Kingdom, and sexual identity. The results are summarized in Figure 1 and Table 2; detailed distributions are shown in Multimedia Appendix 5 and the ORs for each variable are shown in Multimedia Appendix 6.

As Figure 1 shows, for both sexes the average absolute OR was much closer to 1 for Natsal-3 than for the Web surveys suggesting that the probability Natsal-3 sample was more representative of the general population. Among the Web surveys, for both sexes, WS-B2 and WS-M1 had the lowest average absolute ORs.

For men, the largest absolute OR was much lower for Natsal-3 (at 1.32) than for the Web surveys (range 3.13-5.23) (Table 2). The pattern was the same for women, although the differences were not as great. The largest absolute OR was for sexual identity across all 4 Web surveys and for women in Natsal-3 (for men in Natsal-3 it was general health).

The evidence that Web surveys with modified quotas performed better than those with basic quotas was mixed. Among the Web surveys, for men WS-M1 had the lowest average absolute OR, but it was only a small improvement on WS-B2, and both WS-B1 and WS-B2 performed better than the other survey with modified quotas (WS-M2). For women, there was less difference in average absolute ORs between the 4 Web surveys, and the basic quota WS-B2 had the lowest average absolute OR.

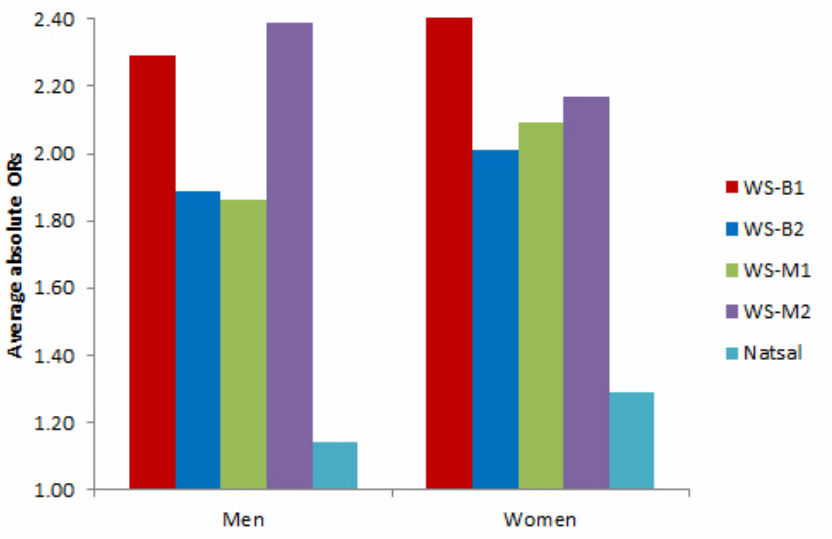
Table 2. Summary of participant characteristics for men and women: Web survey basic quota 1 and 2 (WS-B1 and WS-B2), Web survey modified quota 1 and 2 (WS-M1 and WS-M2), and Natsal-3 compared with independent benchmarks.[a]

Summary of patient characteristics compared with benchmarks	WS-B1	WS-B2	WS-M1	WS-M2	Natsal-3
Men					
Average absolute OR (from benchmarks)	2.29	1.89	1.86	2.39	1.14
Largest absolute OR	5.23	3.45	3.13	3.66	1.32
N significantly different from benchmarks	3/6	5/6	4/6	5/6	2/6
Women					
Average absolute OR (from benchmarks)	2.42	2.01	2.09	2.17	1.29
Largest absolute OR	4.78	3.94	3.83	3.77	1.91
N significantly different from benchmarks	4/6	5/6	6/6	5/6	4/6

[a]The sources of the independent benchmarks were Census 2011 for ethnicity, current economic activity, and self-assessed general health; IHS 2011 for housing tenure and sexual identity; and National Travel Survey 2010 for current driving licence valid in the United Kingdom.

XSL•FO
RenderX

Figure 1. Average absolute odds ratios for participant characteristics for men and women: Web survey basic quota 1 and 2 (WS-B1 and WS-B2), Web survey modified quota 1 and 2 (WS-M1 and WS-M2), and Natsal-3 compared with independent benchmarks.



Comparing Key Estimates from the Web Surveys With Natsal-3 Benchmarks

Comparisons between the Web surveys and Natsal-3 benchmarks for key estimates are summarized in Table 3 by mode (ie, whether the Natsal-3 question was asked in CAPI or CASI) and question type (behavior or opinion) and are shown graphically in Figures 2 and 3. All 4 Web surveys were substantially different from the Natsal-3 benchmarks, as shown in Multimedia Appendix 3 for behaviors and in Multimedia Appendix 7 for opinions, and the ORs for each variable are shown in Multimedia Appendix 8.

The average absolute ORs ranged from 1.32 (WS-B2 women) to 1.98 (WS-B1 men). In all Web surveys, the average absolute ORs were lower for women than for men. Several individual estimates showed very large differences, especially for men; for example, the percentage of men who reported having 1 or more same-sex partners in the past 5 years was 3.0% in Natsal-3 compared with a range of 7.9% to 12.9% in the Web surveys (with ORs ranging from 2.81 to 4.85). Moreover, for each Web survey, a majority of the variables examined were significantly different from the Natsal-3 benchmarks (Figures 4 and 5). The highest percentage of differences between the Web surveys and Natsal-3 was found for the opinion questions (primarily 5-point scales). In 2 of the Web surveys, all 8 opinion questions were significantly different from Natsal-3.

As with sociodemographic characteristics, the evidence that Web surveys with modified quotas performed better than those with basic quotas was mixed. Compared with WS-B1, both WS-M1 and WS-M2 had results (for average absolute ORs) closer to Natsal-3 for both sexes. Compared with WS-B2, however, WS-M1 results showed larger differences in average absolute ORs than WS-B2 for women (but not for men). And, for both sexes, WS-B2 had lower average absolute ORs than did WS-M2. The evidence remains mixed when looking at the different question types and survey modes; for example, WS-M1 had the lowest average absolute ORs for the behavior (whether CAPI or CASI) questions for men, but not for women. Overall, there was not one Web survey that consistently performed better than the others across sex, survey mode, or question type.

Performing tests to compare results from the combined basic quota Web surveys with each modified quota Web survey revealed that the only statistically significant improvement overall (reduction in average absolute OR) was for men in WS-M1 (Table 4).

Another important concern is not only how results from Web panel surveys compare with those using other modes of data collection, but also whether the results obtained are likely to vary significantly according to which Web panel is used. When the 2 basic quota Web surveys were compared using generalized estimating equations, the results were significantly different (P<.001) and differences were often appreciable, as can be seen in in Multimedia Appendix 3.

XSL•FO
RenderX

Table 3. Summary of behavior (CAPI/CASI) and opinion (CAPI/CASI) variables for Web survey basic quota 1 and 2 (WS-B1 and WS-B2) and Web survey modified quota 1 and 2 (WS-M1 and WS-M2) compared with Natsal-3 for men and women.

Summary of behavior and opinion variables compared with Natsal-3	WS-B1	WS-B2	WS-M1	WS-M2
Men, average absolute OR (SE) [a]				
All	1.98 (0.06)	1.62 (0.04)	1.57 (0.03)	1.70 (0.05)
Behavior CAPI	2.24 (0.08)	1.78 (0.05)	1.71 (0.06)	1.83 (0.06)
Behavior CASI	1.80 (0.07)	1.59 (0.05)	1.50 (0.04)	1.61 (0.07)
Opinion CAPI	2.18 (0.11)	1.47 (0.06)	1.43 (0.06)	1.72 (0.09)
Opinion CASI	1.95 (0.10)	1.83 (0.09)	1.79 (0.10)	1.81 (0.10)
Men, largest absolute OR (SE) [a]				
All	4.85 (0.45)	3.28 (0.38)	2.89 (0.28)	3.78 (0.42)
Behavior CAPI	4.54 (0.35)	3.11 (0.24)	2.87 (0.25)	3.75 (0.28)
Behavior CASI	4.85 (0.45)	3.28 (0.38)	2.89 (0.28)	3.78 (0.42)
Opinion CAPI	2.58 (0.15)	2.19 (0.16)	2.28 (0.17)	2.15 (0.16)
Opinion CASI	2.54 (0.15)	2.18 (0.17)	2.24 (0.18)	2.12 (0.17)
Men, significantly different from Natsal-3, n/N				
All	26/35	25/35	26/35	25/35
Behavior CAPI	8/9	7/9	7/9	6/9
Behavior CASI	10/18	12/18	13/18	11/18
Opinion CAPI	5/5	4/5	3/5	5/5
Opinion CASI	3/3	2/3	3/3	3/3
Women, average absolute OR (SE) [a]				
All	1.51 (0.03)	1.32 (0.02)	1.40 (0.03)	1.44 (0.03)
Behavior CAPI	1.54 (0.04)	1.32 (0.04)	1.42 (0.05)	1.59 (0.05)
Behavior CASI	1.32 (0.04)	1.26 (0.03)	1.31 (0.03)	1.28 (0.03)
Opinion CAPI	1.89 (0.13)	1.43 (0.07)	1.42 (0.06)	1.43 (0.06)
Opinion CASI	1.97 (0.11)	1.72 (0.09)	1.84 (0.11)	1.95 (0.11)
Women, largest absolute OR (SE) [a]				
All	2.70 (0.14)	2.04 (0.16)	2.09 (0.14)	2.65 (0.16)
Behavior CAPI	2.70 (0.14)	2.04 (0.16)	2.09 (0.14)	2.65 (0.16)
Behavior CASI	1.91 (0.16)	1.69 (0.18)	1.88 (0.17)	1.82 (0.16)
Opinion CAPI	2.44 (0.34)	1.93 (0.13)	1.92 (0.20)	1.85 (0.17)
Opinion CASI	2.02 (0.17)	1.79 (0.15)	1.91 (0.15)	2.20 (0.16)
Women, significantly different from Natsal-3, n/N				
All	26/35	18/35	24/35	25/35
Behavior CAPI	7/9	4/9	6/9	7/9
Behavior CASI	11/18	8/18	10/18	10/18
Opinion CAPI	5/5	3/5	3/5	5/5
Opinion CASI	3/3	3/3	3/3	3/3

[a] Reference Natsal-3.

XSL•FO
RenderX

Figure 2. Men: Summary of average absolute ORs for behavior (CAPI and CASI) and opinion variables for the 4 Web surveys compared with Natsal-3.

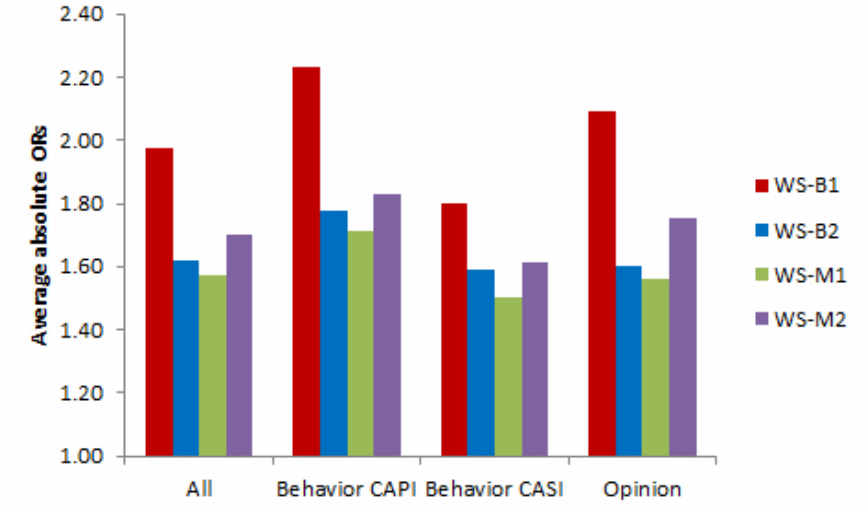


Figure 3. Women: Summary of average absolute ORs for behavior (CAPI and CASI) and opinion variables for the 4 Web surveys compared with Natsal-3.

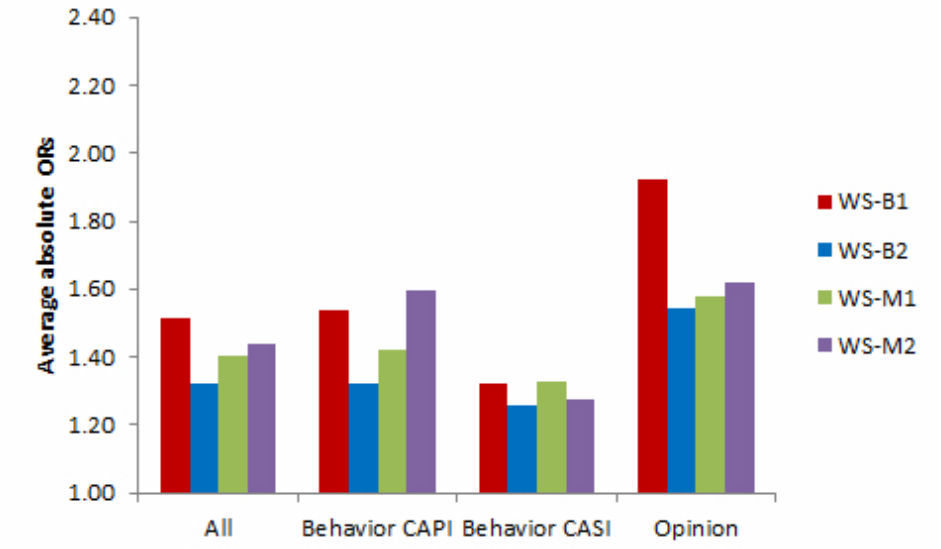
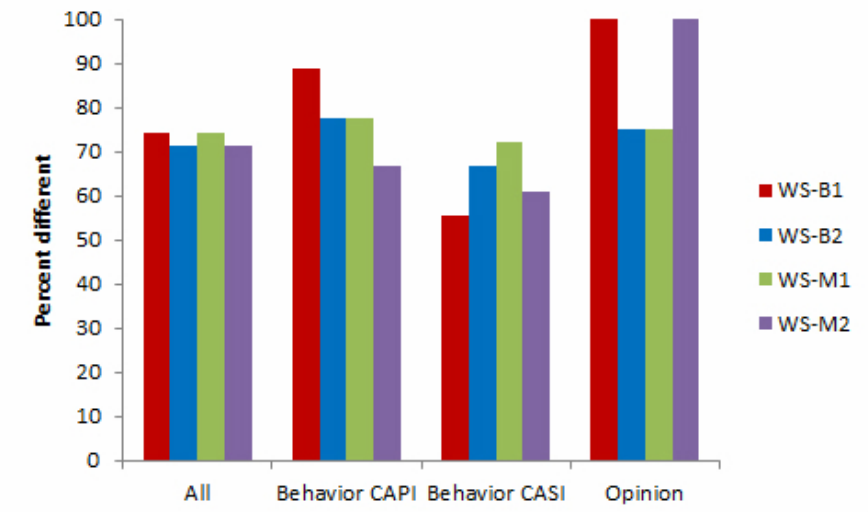


Figure 4. Men: Percentage of variables (all: n=35; behavior CAPI: n=9; behavior CASI: n=18; opinion: n=8) in 4 Web surveys significantly different from Natsal-3.



XSL•FO
RenderX

Figure 5. Women: Percentage of variables (all: n=35; behavior CAPI: n=9; behavior CASI: n=18; opinion: n=8) in 4 Web surveys significantly different from Natsal-3.

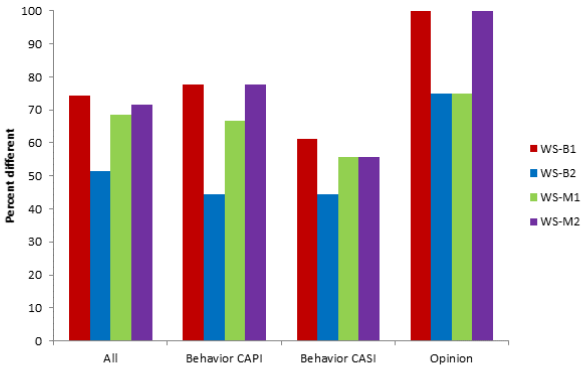


Table 4. Comparison of combined basic quota (WS-B1/WS-B2) Web surveys with first (WS-M1) and second (WS-M2) modified quota Web surveys for behavior and opinion (CAPI/CASI) variables for men and women.

Comparison of basic quota with modified quota Web surveys	WS-B1/WS-B2 combined	WS-M1	*P* [a]	WS-M2	*P* [a]
Men, average absolute OR [b] (SE)					
All	1.80 (0.04)	1.57 (0.03)	<.001	1.70 (0.05)	.10
Behavior CAPI	2.01 (0.05)	1.71 (0.06)	<.001	1.83 (0.06)	.02
Behavior CASI	1.70 (0.05)	1.50 (0.04)	<.001	1.61 (0.07)	.26
Opinion CAPI	1.83 (0.07)	1.43 (0.06)	<.001	1.72 (0.09)	.31
Opinion CASI	1.89 (0.07)	1.79 (0.10)	.40	1.81 (0.10)	.52
Women, average absolute OR [b] (SE)					
All	1.42 (0.02)	1.40 (0.03)	.54	1.44 (0.03)	.55
Behavior CAPI	1.43 (0.03)	1.42 (0.05)	.85	1.59 (0.05)	.01
Behavior CASI	1.29 (0.02)	1.31 (0.03)	.62	1.28 (0.03)	.79
Opinion CAPI	1.66 (0.08)	1.42 (0.06)	.01	1.43 (0.06)	.01
Opinion CASI	1.85 (0.07)	1.84 (0.11)	.94	1.95 (0.11)	.45

[a] *P* values were calculated using the bootstrap method as described in the Methods section.

[b] From benchmarks.

Discussion

As the first study in Britain to compare surveys using volunteer Web panels maintained by different research agencies, we demonstrate that, as in other countries, they are not reliably able to provide scientifically robust results. Between 60% and 75% of key estimates in each of 4 Web surveys were significantly different from the benchmarks provided by Natsal-3, a high-quality probability sample survey. Differences between the Web surveys and Natsal-3 were very large for some key behaviors, such as estimates of same-sex sexual experience, which were 2 to 3 times higher in the Web surveys. There were important differences in sample composition as well; for example, compared with both Natsal-3 and the general population, the Web surveys contained more men and women who self-identified as gay/bisexual and fewer men and women from nonwhite ethnic groups.

One limitation of this study is that it is not possible to determine the extent to which the differences observed between the Web surveys and Natsal-3 were due to differences in sample

XSL•FO
RenderX

composition or to differences in mode of data collection. Another limitation is that, for the key estimates of sexual behavior, it is not possible to say definitively whether Natsal-3 or the Web surveys provide the most "accurate" estimates. However, we view Natsal-3 as likely to be less biased given the comparisons with independent benchmarks, which showed that participants in the Web panel surveys were less representative of the general population on a number of sociodemographic and other characteristics than Natsal-3 participants. This was found even for the relatively young population (18 to 44 years) included in this study. This is an age group which does not suffer from undercoverage due to lack of Internet use or access because more than 90% of the 18-44 years age group in Britain use the Internet at least once a week and live in households with Internet access [34]. Differences in sample composition, therefore, are likely to explain at least some of the differences observed in the behavioral estimates between the Web surveys and Natsal-3. Further, our findings suggest that setting quotas on key demographics (eg, age), or even on variables known to be correlated to key estimates, are unlikely to ensure representativeness on demographic (or other) variables that are not used to set quotas.

The average absolute ORs in the Web surveys were higher for the Natsal-3 behavior variables asked in CAPI than in CASI. This is consistent with the possibility that Web surveys, because of the greater anonymity afforded by the Web, may obtain higher rates of disclosure of sensitive behaviors than a CAPI survey, and that fewer differences would be expected between a Web survey and sensitive questions asked in CASI. However, even though both CASI and Web surveys are self-administered, they differ in other respects, such as the presence of an interviewer and the degree of perceived privacy. Although the Web surveys obtained much higher rates of disclosure of same-sex attraction and experience than obtained in either CAPI or CASI in Natsal-3, over a range of variables there was no consistent pattern found in whether reports of sensitive behaviors were higher in the Web surveys or in CASI. For some variables, there were only small differences, or indeed lower reporting, in the Web surveys than in Natsal-3 CASI. We found that approximately one-half to two-thirds of the Web survey estimates were significantly different from the Natsal-3 CASI estimates. Because the visually presented self-administered modes (ie, Web and CASI) do not yield similar estimates for the majority of questions, it appears that differences in mode (ie, measurement error) cannot fully explain the differences in the estimates between the Web surveys and Natsal-3, and that selection biases must also be present.

Of the 4 Web surveys, it was not possible to identify one that performed consistently better than the others, either over all estimates or in groups defined by survey mode in Natsal-3 (CAPI vs CASI) or question type (behavior vs opinions). This lack of predictability in how the Web surveys compare across estimates is consistent with findings for less sensitive measures in the United States [15,16].

Two of the Web surveys included additional quota controls known to be related to key estimates, but this did not lead to any consistent improvement. Although the modified quota Web surveys showed improved estimates for some variables, others did not change and yet others moved further away from the benchmarks. Moreover, the modified quotas presented operational difficulties for the 2 agencies carrying out the Web surveys and had to be relaxed in both cases.

Consistent with findings reported elsewhere that participants answering opinion questions in a self-administered mode (eg, on the Web or in CASI) are more likely to "satisfice" (ie, make less cognitive effort to think about and answer a survey question) than are those to a personal interview [13,19,35,36], the Web surveys and Natsal-3 CASI showed much greater use of neutral points (ie, "don't know," "depends," or "neither agree nor disagree") when compared with the same (or similar) opinion questions in Natsal-3 CAPI (Multimedia Appendix 7).

Although only a shortened version of the Natsal-3 questionnaire was included in the Web surveys, the questions were included in the same order as in Natsal-3 and were identically worded. Many of the questions excluded from the Web surveys had to do with the participant's health condition and were aimed at the older age group (45-74 years) in Natsal-3 but who were not included in the Web survey sample. Also, in this paper, we have only looked at overall differences in prevalence estimates for men and women, and not at any other subgroup analysis and we have not looked at relationships among variables, but this analysis is planned for a subsequent paper. In further work, we shall also examine the degree to which other weighting/adjustment can reduce the bias seen here in the Web surveys.

Finally, we were not able to compare Natsal-3 results with those from a Web panel selected using probability sampling methods because no such panels have been recruited in Britain. An interesting future study would be to compare results between Natsal-3 and an Internet survey recruited using probability sampling methods, although the limited evidence available elsewhere suggests that the Internet survey results are still likely to differ significantly from those using face-to-face interviewing methods [37].

The demand for Web surveys, including in the field of health and epidemiological surveys, is likely to continue to increase as researchers look for new methods of data collection that are cost-effective while maintaining scientific rigor [38]. Commissioning Web surveys among the volunteer panels maintained by large market research agencies is a possible route because they are able to provide a quick turnaround at a much lower cost than traditional interview methods. Although using these volunteer Web panels may be suitable for certain types of study—potentially including surveys of hard-to-reach groups [4], for testing the properties of psychometric questionnaires [14], for syndromic surveillance [39], or for epidemiological studies where representative sampling may not be required [40]—the evidence from our investigation within Britain supports the conclusion that such surveys are not appropriate substitutes for probability-based sample surveys that aim to provide scientifically robust population estimates.

XSL•FO
RenderX

Acknowledgments

Natsal-3 was funded by grants from the UK Medical Research Council (G0701757) and the Wellcome Trust (084840), with support from the Economic and Social Research Council and the Department of Health. The 4 Web panel surveys were funded by a supplementary grant from the Wellcome Trust. The sponsors played no role in the study design, data interpretation, data analysis, or writing of the paper. Natsal-3 is a collaboration between University College London (London, UK), the London School of Hygiene & Tropical Medicine (London, UK), NatCen Social Research (London, UK), Public Health England (formerly the Health Protection Agency), and the University of Manchester (Manchester, UK). The supplementary study of Web panel surveys also included collaborators from the University of Michigan (Ann Arbor, MI, USA). We thank the study participants in Natsal-3 and in the Web panel surveys, the team of interviewers at NatCen Social Research, and the researchers and programmers at the 3 market research agencies that carried out the Web panels.

Authors' Contributions

The paper was conceived by BE, AJC, SB, MPC and FC. BE wrote the first draft, with further contributions from SB, AJC, MPC, FC, SC, CT, AP, JD, CHM, PS, PP, KRM, KW and AMJ. SB carried out the statistical analysis with support from AJC. BE, AJC, AMJ, MPC and FC wrote the protocol and obtained funding for this methodological sub-study of Natsal-3. BE, AJC, CHM, PS, KW and AMJ (Principal Investigator) were initial applicants, wrote the study protocol and obtained funding for Natsal-3. BE, SC and AP designed the web version of the questionnaire, and they, along with CT, JD, CHM, PS, KRM, KW and AMJ designed the Natsal-3 questionnaire and applied for ethics approval. All authors contributed to data interpretation, reviewed successive drafts and approved the final version of the manuscript.

Conflicts of Interest

Anne Johnson has been a Governor of the Wellcome Trust since 2011. The other authors declare that they have no conflicts of interest.

Multimedia Appendix 1

Natsal-3 questionnaire content.

[PDF File (Adobe PDF File), 27KB - jmir_v16i12e276_app1.pdf]

Multimedia Appendix 2

Natsal Web questionnaire.

[PDF File (Adobe PDF File), 294KB - jmir_v16i12e276_app2.pdf]

Multimedia Appendix 3

Behavioral variables: 4 Web surveys compared with Natsal-3, separately for men and women.

[PDF File (Adobe PDF File), 293KB - jmir_v16i12e276_app3.pdf]

Multimedia Appendix 4

Achieved and target quotas for the modified quota Web surveys.

[PDF File (Adobe PDF File), 150KB - jmir_v16i12e276_app4.pdf]

Multimedia Appendix 5

Participant characteristics: 4 Web surveys and Natsal-3 compared with independent benchmarks, separately for men and women.

[PDF File (Adobe PDF File), 82KB - jmir_v16i12e276_app5.pdf]

Multimedia Appendix 6

ORs (95% CIs) for participant characteristics: 4 Web surveys and Natsal-3 compared with independent benchmarks, separately for men and women.

[PDF File (Adobe PDF File), 16KB - jmir_v16i12e276_app6.pdf]

XSL•FO
RenderX

Multimedia Appendix 7

Opinion variables: 4 Web surveys compared with Natsal-3, separately for men and women.

[PDF File (Adobe PDF File), 174KB - jmir_v16i12e276_app7.pdf]

Multimedia Appendix 8

ORs (95% CIs) for behavioral and opinion variables: 4 Web surveys compared with Natsal-3, separately for men and women.

[PDF File (Adobe PDF File), 214KB - jmir_v16i12e276_app8.pdf]

References

1. Touvier M, Méjean C, Kesse-Guyot E, Pollet C, Malon A, Castetbon K, et al. Comparison between web-based and paper versions of a self-administered anthropometric questionnaire. Eur J Epidemiol 2010 May;25(5):287-296. [doi: 10.1007/s10654-010-9433-9] [Medline: 20191377]
2. Shah AD, Wood DM, Dargan PI. Internet survey of home storage of paracetamol by individuals in the UK. Q J Med 2013;106(3):253-259 [FREE Full text] [doi: 10.1093/qjmed/hcs232 253-259]
3. Charrance G, Legleye S, Razafindrastsima N. The use of an internet panel to monitor sexual and reproductive health in the general population. 2013. URL: http://www.europeansurveyresearch.org/conf/uploads/466/695/184/CHARRANCE_ONLINE_SURVEYS.pdf? [accessed 2014-03-03] [WebCite Cache ID 6Nnfhcawf]
4. Marcus U, Hickson F, Weatherburn P, Schmidt AJ, EMIS Network. Prevalence of HIV among MSM in Europe: comparison of self-reported diagnoses from a large scale internet survey and existing national estimates. BMC Public Health 2012;12:978 [FREE Full text] [doi: 10.1186/1471-2458-12-978] [Medline: 23151263]
5. Couper MP, Singer E, Conrad FG, Groves RM. Experimental Studies of Disclosure Risk, Disclosure Harm, Topic Sensitivity, and Survey Participation. J Off Stat 2010 Jun;26(2):287-300 [FREE Full text] [Medline: 21765576]
6. Ekman A, Litton JE. New times, new needs; e-epidemiology. Eur J Epidemiol 2007;22(5):285-292. [doi: 10.1007/s10654-007-9119-0] [Medline: 17505896]
7. Ekman A, Dickman PW, Klint A, Weiderpass E, Litton JE. Feasibility of using web-based questionnaires in large population-based epidemiological studies. Eur J Epidemiol 2006;21(2):103-111. [doi: 10.1007/s10654-005-6030-4] [Medline: 16518678]
8. Link MW, Mokdad AH. Alternative modes for health surveillance surveys: an experiment with web, mail, and telephone. Epidemiology 2005 Sep;16(5):701-704. [Medline: 16135951]
9. Hou SI, Wisenbaker J. Using a web-based survey to assess correlates of intention towards HIV testing among never-been-tested but sexually experienced college students. AIDS Care 2005 Apr;17(3):329-334. [doi: 10.1080/09540120412331299816] [Medline: 15832880]
10. Link MW, Mokdad AH. Are web and mail modes feasible options for the behavioral risk factor surveillance system? In: Cohen SB, Lepkowski JM, editors. Eighth Conference on Health Survey Research Methods. 2004 Presented at: Eighth Conference on Health Survey Research Methods. Natl Cent Health Statistics; February 20-23, 2004; Peachtree City, GA URL: http://www.cdc.gov/nchs/data/hsrmc/hsrmc_8th_proceedings_2004.pdf
11. Baker R, Zahs D, Popa G. Health surveys in the 21st century: telephone vs web. In: Cohen SB, Lepkowski JM, editors. Eighth Conference on Health Survey Research Methods. 2004 Presented at: Eighth Conference on Health Survey Research Methods. Natl Cent Health Statistics; February 20-23, 2004; Peachtree, GA URL: http://www.cdc.gov/nchs/data/hsrmc/hsrmc_8th_proceedings_2004.pdf
12. Baer A, Saroiu S, Koutsky LA. Obtaining sensitive data through the Web: an example of design and methods. Epidemiology 2002 Nov;13(6):640-645. [doi: 10.1097/01.EDE.0000032360.92664.BC] [Medline: 12410004]
13. AAPOR Report on Online Panels. 2010 Mar. URL: http://www.aapor.org/AM/Template.cfm?Section=AAPOR_Committee_and_Task_Force_Reports&Template=/CM/ContentDisplay.cfm&ContentID=2223 [WebCite Cache ID 6Nnji0q7q]
14. Craig BM, Hays RD, Pickard AS, Cella D, Revicki DA, Reeve BB. Comparison of US panel vendors for online surveys. J Med Internet Res 2013;15(11):e260 [FREE Full text] [doi: 10.2196/jmir.2903] [Medline: 24292159]
15. Yeager DS, Krosnick JA, Chang L, Javitz HS, Levendusky MS, Simpser A, et al. Comparing the accuracy of RDD telephone surveys and internet surveys conducted with probability and non-probability samples. Public Opin Q 2011;75(4):709-747. [doi: 10.1093/poqnfr020]
16. Chang LC, Krosnick JA. National surveys via RDD telephone interviewing versus the internet: comparing sample representativeness and response quality. Public Opin Q 2009;73(4):641-678. [doi: 10.1093/poqnfp075]
17. Loosveldt G, Sonck N. An evaluation of the weighting procedures for an online access panel survey. Surv Res Methods 2008;2(2):93-105 [FREE Full text]
18. Smith P. SRA News. 2013. Surveying the UK general population through the web URL: http://the-sra.org.uk/wp-content/uploads/sra_research_matters_march_2013.pdf [accessed 2014-03-05] [WebCite Cache ID 6NqU4qGzI]

XSL•FO
RenderX

19. Duffy B, Smith K, Terhanian G, Bremer J. Comparing data from online and face-to-face surveys. Int J Mark Res 2005;47(6):615-639 [FREE Full text]
20. Gosling SD, Vazire S, Srivastava S, John OP. Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. Am Psychol 2004;59(2):93-104. [doi: 10.1037/0003-066X.59.2.93] [Medline: 14992636]
21. Cabinet Office. Community Life Survey: Development of content and methodology for future survey years. URL: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/325868/Community_Life_Survey_consultation_-_November_2013__3_.pdf [accessed 2014-03-05] [WebCite Cache ID 6UBIfNW1Y]
22. Dayan Y, Paine CS, Johnson AJ. Responding to sensitive questions in surveys: a comparison with results from Online panels, face to face and self-completion interviews. URL: http://www.ipsos-mori.com/Assets/Docs/Publications/Ops_RMC_Responding_Sensitive_Questions_08_03_10.pdf [accessed 2014-03-03] [WebCite Cache ID 6NnhbsTSm]
23. Kreuter F, Presser S, Tourangeau R. Social desirability bias in CATI, IVR, and Web surveys: the effects of mode and question sensitivity. Public Opinion Quarterly 2008;72(5):847-865. [doi: 10.1093/poq/nfn063]
24. Erens B, Phelps A, Clifton S, Mercer CH, Tanton C, Hussey D, et al. Methodology of the third British National Survey of Sexual Attitudes and Lifestyles (Natsal-3). Sex Transm Infect 2014;90(2):84-89 [FREE Full text] [doi: 10.1136/sextrans-2013-051359] [Medline: 24277881]
25. Office for National Statistics. Labour Force Survey User Guide 2009-Volume 1: Background and Methodology. 2009. URL: http://www.ons.gov.uk/ons/guide-method/method-quality/specific/labour-market/labour-market-statistics/index.html [accessed 2014-03-03] [WebCite Cache ID 6NnglvlaO]
26. Payne RA, Abel GA. UK indices of multiple deprivation - a way to make comparisons across constituent countries easier. Health Stat Q 2012(53):22-37. [Medline: 22410690]
27. Office for National Statistics. Integrated Household Survey User Guide-Volume 1: IHS Background and Methodology. 2012. URL: http://www.ons.gov.uk/ons/guide-method/method-quality/specific/social-and-welfare-methodology/integrated-household-survey/index.html [accessed 2014-03-03] [WebCite Cache ID 6Nnh5cbFC]
28. Rofique J, Humphrey A, Pickering K, Tipping S. National Travel Survey 2010 Technical Report. 2011 Jun. URL: http://doc.ukdataservice.ac.uk/doc/5340/mrdoc/pdf/5340_technical_report_nts2010.pdf [accessed 2014-03-03] [WebCite Cache ID 6UUOEEBkl]
29. Mercer CH, Tanton C, Prah P, Erens B, Sonnenberg P, Clifton S, et al. Changes in sexual attitudes and lifestyles in Britain through the life course and over time: findings from the National Surveys of Sexual Attitudes and Lifestyles (Natsal). Lancet 2013 Nov 30;382(9907):1781-1794 [FREE Full text] [doi: 10.1016/S0140-6736(13)62035-8] [Medline: 24286784]
30. Sonnenberg P, Clifton S, Beddows S, Field N, Soldan K, Tanton C, et al. Prevalence, risk factors, and uptake of interventions for sexually transmitted infections in Britain: findings from the National Surveys of Sexual Attitudes and Lifestyles (Natsal). Lancet 2013 Nov 30;382(9907):1795-1806 [FREE Full text] [doi: 10.1016/S0140-6736(13)61947-9] [Medline: 24286785]
31. Wellings K, Jones KG, Mercer CH, Tanton C, Clifton S, Datta J, et al. The prevalence of unplanned pregnancy and associated factors in Britain: findings from the third National Survey of Sexual Attitudes and Lifestyles (Natsal-3). Lancet 2013 Nov 30;382(9907):1807-1816 [FREE Full text] [doi: 10.1016/S0140-6736(13)62071-1] [Medline: 24286786]
32. Mitchell KR, Mercer CH, Ploubidis GB, Jones KG, Datta J, Field N, et al. Sexual function in Britain: findings from the third National Survey of Sexual Attitudes and Lifestyles (Natsal-3). Lancet 2013 Nov 30;382(9907):1817-1829 [FREE Full text] [doi: 10.1016/S0140-6736(13)62366-1] [Medline: 24286787]
33. Efron B, Tibshirani R. An Introduction to the Bootstrap. New York: Chapman & Hall; 1993.
34. Office for National Statistics. Statistical Bulletin. 2012. Internet Access-Households and Individuals 2012 URL: http://www.ons.gov.uk/ons/dcp171778_275775.pdf [accessed 2014-03-04] [WebCite Cache ID 6NooqeMk3]
35. Chang L, Krosnick JA. Comparing oral interviewing with self-administered computerized questionnaires. An experiment. Public Opin Q 2010;74(1):154-167. [doi: 10.1093/nfp090]
36. Heerwegh D, Loosveldt G. Face-to-face versus Web surveying in a high-Internet-coverage population: differences in response quality. Public Opin Q 2008;72(5):836-846. [doi: 10.1093/poq/nfn045]
37. Cabinet Office. Community Life Survey: Summary of web experiment findings. 2013 Nov. URL: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/325872/Annex_B_-_Summary_of_web_experiment_findings_2012-13.pdf [accessed 2014-07-28] [WebCite Cache ID 6RP8yP1BH]
38. Schonlau M. Will web surveys ever become part of mainstream research? J Med Internet Res 2004 Sep 23;6(3):e31 [FREE Full text] [doi: 10.2196/jmir.6.3.e31] [Medline: 15471757]
39. Sugiura H, Ohkusa Y, Akahane M, Sano T, Okabe N, Imamura T. Development of a web-based survey for monitoring daily health and its application in an epidemiological survey. J Med Internet Res 2011;13(3):e66 [FREE Full text] [doi: 10.2196/jmir.1872] [Medline: 21946004]
40. Richiardi L, Pizzi C, Pearce N. Commentary: Representativeness is usually not necessary and often should be avoided. Int J Epidemiol 2013 Aug;42(4):1018-1022 [FREE Full text] [doi: 10.1093/ije/dyt103] [Medline: 24062290]

XSL•FO
RenderX

Abbreviations

CAPI: computer-assisted personal interview
CASI: computer-assisted self-interview
IHS: Integrated Household Survey
IMD: Index of Multiple Deprivation
LISS: Longitudinal Internet Study for the Social sciences
Natsal-3: Third National Survey of Sexual Attitudes and Lifestyles
ONS: Office for National Statistics

Edited by G Eysenbach; submitted 07.03.14; peer-reviewed by P Lynn, C Turner, S Legleye; comments to author 16.07.14; revised version received 02.09.14; accepted 06.09.14; published 08.12.14.

Please cite as:
Erens B, Burkill S, Couper MP, Conrad F, Clifton S, Tanton C, Phelps A, Datta J, Mercer CH, Sonnenberg P, Prah P, Mitchell KR, Wellings K, Johnson AM, Copas AJ
Nonprobability Web Surveys to Measure Sexual Behaviors and Attitudes in the General Population: A Comparison With a Probability Sample Interview Survey
J Med Internet Res 2014;16(12):e276
URL: http://www.jmir.org/2014/12/e276/
doi:10.2196/jmir.3382
PMID:25488851

©Bob Erens, Sarah Burkill, Mick P Couper, Frederick Conrad, Soazig Clifton, Clare Tanton, Andrew Phelps, Jessica Datta, Catherine H Mercer, Pam Sonnenberg, Philip Prah, Kirstin R Mitchell, Kaye Wellings, Anne M Johnson, Andrew J Copas. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 08.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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

Recruitment via the Internet and Social Networking Sites: The 1989-1995 Cohort of the Australian Longitudinal Study on Women's Health

Gita Devi Mishra[1], CScient, CStat, PhD; Richard Hockey[1], BSC; Jennifer Powers[2], MMedStat; Deborah Loxton[2], PhD; Leigh Tooth[1], PhD; Ingrid Rowlands[1], PhD; Julie Byles[2], PhD; Annette Dobson[1], PhD

[1]Centre for Longitudinal and Life Course Research, School of Public Health, University of Queensland, Herston, Australia

[2]Priority Research Centre for Gender, Health and Ageing, University of Newcastle, Newcastle, Australia

Corresponding Author:
Gita Devi Mishra, CScient, CStat, PhD
Centre for Longitudinal and Life Course Research
School of Public Health
University of Queensland
Herston Road
Herston, 4006
Australia
Phone: 61 7 33465224
Fax: 61 7 33655540
Email: g.mishra@uq.edu.au

Abstract

Background: Faced with the challenge of recruiting young adults for health studies, researchers have increasingly turned to the Internet and social networking sites, such as Facebook, as part of their recruitment strategy. As yet, few large-scale studies are available that report on the characteristics and representativeness of the sample obtained from such recruitment methods.

Objective: The intent of the study was to describe the sociodemographic and health characteristics of a national sample of young Australian women recruited mainly through the Internet and social networking sites and to discuss the representativeness of their sociodemographic, health, and lifestyle characteristics relative to the population.

Methods: A cohort of 17,069 women (born between 1989 and 1995) was recruited in 2012-13 for the Australian Longitudinal Study on Women's Health. Sociodemographic characteristics (percentages, means, and 95% confidence intervals) from the online survey data were compared with women aged 18-23 years from the 2011 Australian Census. Sample data were compared by age and education level with data from the 2011-13 Australian Health Survey (AHS).

Results: Compared to the Australian Census data, study participants were broadly representative in terms of geographical distribution across Australia, marital status (95.62%, 16,321/17,069) were never married), and age distribution. A higher percentage had attained university (22.52%, 3844/17,069) and trade/certificate/diploma qualifications (25.94%, 4428/17,069) compared with this age group of women in the national population (9.4% and 21.7% respectively). Among study participants, 22.05% (3721/16,877) were not in paid employment with 35.18% (5931/16,857) studying 16 or more hours a week. A higher percentage of study participants rated their health in the online survey as fair or poor (rather than good, very good, or excellent) compared with those participating in face-to-face interviews in the AHS (18.77%, 3203/17,069 vs 10.1%). A higher percentage of study participants were current smokers (21.78%, 3718/17,069 vs 16.4%) and physically active (59.30%, 10,089/17,014 were classified as sufficiently active vs 48.3%) but alcohol consumption was lower (59.58%, 9865/16,558 reported drinking alcohol at least once per month vs 65.9% in the AHS). Using self-reported height and weight to determine body mass index (BMI, kg/m^2), 34.80% (5901/16,956) of the cohort were classified as overweight or obese (BMI of 25 or more), compared with 33.6% respectively using measured height and weight in the AHS.

Conclusions: Findings indicated that using the Internet and social networking sites for an online survey represent a feasible recruitment strategy for a national cohort of young women and result in a broadly representative sample of the Australian population.

XSL•FO
RenderX

(J Med Internet Res 2014;16(12):e279) doi:10.2196/jmir.3788

KEYWORDS

Web-based survey; online survey; social media; Facebook; representativeness; education; socioeconomic factors; women's health; data collection; young adults; Australia

Introduction

Recruitment of participants for longitudinal health studies poses increasing challenges for researchers, with indications of declining participation rates in telephone- or mail-based data collection surveys [1-3]. Recruitment and retention are particularly acute issues with respect to young adults, who are highly mobile and difficult to reach by conventional modes of contact, such as landline telephone or postal address [4,5]. Young adults' familiarity with digital and mobile technologies, however, provides an opportunity for innovative recruitment and survey modalities including the Internet and social networks. Although recent research suggests that participant recruitment via social networking sites has advantages over traditional methods (eg, convenient, cost effective, reaches young adults), it is often described as introducing a participant self-selection bias, limiting the representativeness of the sample [6]. The issue of sample representativeness is the subject of ongoing debate but its relevance requires consideration of the research questions and study design [7-14]. Representativeness may not be important, or even desirable, for etiological studies, if the measurement and control of confounders is feasible [9,12]. However, representativeness is important if one of the goals of the study is to estimate the prevalence of disease or health status of population groups [11,13]. Also, having a sample of sufficient size and diversity in terms of a range of key characteristics and exposures is an essential attribute for many epidemiological studies [3,8,15]. It remains unclear, however, if recruitment strategies that use the Internet and social networks are able to obtain a representative sample of young adults for large national cohort studies.

Given the logistics and financial challenges of using conventional survey methods, increasingly the Internet, in conjunction with integrated database management systems, has been used to access a large sampling frame of potential participants [16-18]. Yet a distinction should be drawn between relying on social network sites, such as Facebook, for recruitment and the separate implementation of an online survey where a variety of recruitment methods direct participants to the survey website. The latter is exemplified by the recent French NutriNet-Santé study where television advertisements provided the major strategy for recruitment, achieving over 88,000 participants, but promotion of the study via the Internet and radio also contributed substantially [19]. Some large established cohort studies, such as the Black Women's Health Study [20], have also changed their data collection methods to online surveys for the follow-up surveys. The Australian Longitudinal Study on Women's Health (ALSWH), which includes three age cohorts, has transferred the two younger cohorts, one born in 1973-1978, the other born in 1946-1951 (originally recruited in 1996), from postal questionnaires to online surveys (with postal paper surveys available, if requested) [21].

While online or Web-based questionnaires can assist with survey completion, they do not specifically address the issue of recruiting a representative sample of young adults. The increased access to the Internet by young adults and the concomitant rise in popularity of online social networks has provided a way forward for health researchers. Recent surveys in the United States have found that 76% and 83% of 18-29 year olds have access to broadband or a smartphone respectively [22,23], and 73% of adults who go online use a social networking site of some kind [24]. Facebook has emerged as the preeminent social networking platform, with an estimated 1.2 billion monthly users and monthly usage statistics indicating 12 million unique Australian visitors [25]. Overall, young Australian adults display similar Internet usage patterns to those evident in the United States [26], with 92% of 18 to 24 year olds visiting social networking sites on a regular basis [27], most (95%) of whom are Facebook members [28].

Social networks can be defined as groups of people with some shared pattern of contacts or interactions between them [29]. Researchers have used Facebook advertising to target people with specific health conditions and lifestyles [4,26,30-33], but these studies have typically been small scale [4,31-33] or rely on an identified attribute that characterizes the network to facilitate a snowball recruitment strategy (eg, peer referral to the study) [34]. The reliance on social connections may be challenging for national health surveys that seek a representative or comprehensive sample of the population [33] and—importantly for the Australian context—a sufficient sample of young adults from rural and remote areas. However, Facebook offers a convenient, immediate, and low-cost way to contact a broad sample of eligible young adults and targets advertising dynamically to specific sociodemographic groups that are under-represented among study participants.

The paper reports on a large national sample of young Australian women (aged 18 to 23 years) and who were primarily recruited through Facebook advertising and other Internet-based modes of contact. We compare the sociodemographic, health, and lifestyle characteristics of this cohort with women in the same age range from the 2011 Australian Census and the 2011-12 Australian Health Survey (AHS).

Methods

Study Design

Since its baseline survey in 1996 of over 40,000 Australian women, the ALSWH has become established as the Australia-wide study of women's health, with surveys conducted approximately every 3 years since 1998 [35]. Until recently, the study comprised three cohorts of women born in 1921-26, 1946-51, and 1973-78. These women were randomly selected

XSL•FO
RenderX

using the national health insurance database (Medicare), which includes all permanent residents of Australia. Comparison of demographic characteristics of participants at baseline with census data indicated that the samples are broadly representative of the Australian population in these age groups [36].

This paper uses data collected from a new young ALSWH cohort of 17,069 women born from 1989-95 and recruited in 2012-13. Women will be surveyed annually with the primary aim of identifying changes in health and well-being and health service needs across the lifespan, to inform Australian policies across a range of issues. Eligible women were those aged 18-23 years when they completed the surveys, who had a valid Medicare number (this includes all permanent residents, but not temporary residents, such as overseas students). The women also needed to consent to having their survey data linked with administrative health data on their health service utilization. Approval for the study was obtained from the Human Research Ethics Committee of the University of Newcastle and the University of Queensland, as well as the Department of Human Services and the Department of Health. Further details of the survey methodology are available from the study website [21].

Recruitment

Initially we planned to recruit the new cohort of young women using the same methodology adopted for the previous cohorts, with contact by mail, however, this approach was reassessed when a pilot survey using these methods for another Australian study with women of a similar age yielded only a 6% response rate [5]. Subsequently, an array of recruitment strategies was deployed: advertising through Facebook or other online media (eg, study website, Gumtree, Twitter, Instagram, Tumblr, YouTube), referral (word-of-mouth by study staff members and their networks, professional bodies, and participants who had already completed the survey), and conventional media advertising (eg, posters, flyers, magazines, TV, and radio interviews). Cinema advertising was also tried in some regional areas. Over the recruitment period, two distinct campaigns were conducted. The first was designed by study staff members and offered the chance to win one of 100 AU $50 gift vouchers (October 2012-September 2013) and the second was coordinated by a marketing company and offered the chance to win one of 2000 exclusive pairs of leggings designed by an independent clothing designer (October 2013-December 2013). Resources progressively shifted from conventional media to online social media according to the observed response rates. Of all the methods adopted, targeted advertising through Facebook was the most successful means of recruitment (69.94%, 11,799/16,869), followed by the marketing company campaign (12.72%, 2145/16,869), referral (7.02% 1184/16,869), conventional media (5.39%, 910/16,869), and other online media (4.93%, 831/16,869).

Data were collected via a Web-based survey. Eligibility was assessed by asking participants to submit their personal and contact details. Eligible study participants were asked 62 questions on: sociodemographic and personal characteristics (eg, educational qualifications), aspects of physical and mental health (eg, self-rated general health), anthropometric data (height, weight), health risk factors (eg, physical activity levels), risk-taking behavior (eg, illicit drug use), access to health service use (eg, screening services), reproductive health and outcomes (eg, pregnancy, birth outcomes), and experience of violence or abuse. Survey features, such as organizing the questions by topic, limiting the number of questions to only one or two per page, using a multiple choice format where possible, and a visible progress bar were used to encourage survey completion and to minimize participant burden.

Demographic data from the study participants were routinely compared with 2011 Census data. The Australian Bureau of Statistics conducts the Australian Census every 5 years, with the most recent being on 9 August 2011. The Census measures key sociodemographic characteristics of all people who are in Australia on Census Night, including their education level and marital status [37]. Advertising strategies were then dynamically adapted according to the areas or demographic groups identified as being under-represented in the sample as it accrued. The recruitment period for the cohort ran for 14 months from 26 October 2012 to 19 December 2013.

Characteristics of the study participants were also compared with women in the same age group from the 2011-13 Australian Health Survey (AHS), a large national health survey. Initially, 30,721 households were approached and of these, 25,080 (81.64%) responded, resulting in 31,837 participants. Face-to-face interviews with one adult from each household collected data on a range of health-related issues, including health status, risk factors, socioeconomic circumstances, physical activity, and nutrition [38]. The height and weight of participants, used for the BMI classification, were obtained from measured rather than self-reported data.

Sociodemographic, Health-Related, and Lifestyle Characteristics

Data for sociodemographic variables were re-categorized to facilitate comparison with the 2011 Australian Census data: age (in years); State/Territory of residence; area of residence based on an index of distance to the nearest urban center (major cities, inner regional, outer regional, remote, very remote) [39]; education level completed (less than year 12, year 12 or equivalent, certificate/diploma, university degree); Aboriginal or Torres Strait Island origin (no, yes); and current relationship status (never married, married, separated/divorced/widowed).

Similarly, data for health-related and lifestyle variables were recoded to enable comparison with the 2011-2013 AHS, as follows: self-rated health (excellent, very good, good, fair or poor); smoking status (non-smoker, current smoker); body weight (kg); height (cm); body mass index (BMI); underweight (<18.5 kg/m^2); normal weight (18.5 to <25 kg/m^2); overweight (25 to <30 kg/m^2); obese (≥30 kg/m^2), according to the World Health Organization's classification [40]; and alcohol consumption (never drink, less than once a month, less than once a week, at least once a week) [41]. A physical activity category was derived from questions on the frequency and duration of different types of physical activity (inactive, insufficiently active, sufficiently active) [42].

XSL•FO
RenderX

Statistical Analysis

The sociodemographic characteristics of the sample (percentages, means, and 95% confidence interval) were compared with corresponding data from women in the same age group in the 2011 Census. The prevalence of health-related and lifestyle characteristics were then compared with the 2011-2013 AHS. Based on preliminary analysis, and to enable comparison with the AHS data, weights for the sample, W(x), at each education level x, were calculated as:

$$W(x) = (N/P) \times (P(x)/N(x))$$

where N is the number of women in the sample and N(x) is the number of women in the sample with education level x. Similarly, P is the number of women in the 18-23 year age group in the Australian population and P(x) is the number of women in the 18-23 year age-group in the Australian population with education level x. Women who had missing data for their education level (7.8%) were omitted from the calculation of weights, which in effect assumes that the data are missing at random. The unweighted and the weighted data are presented.

Results

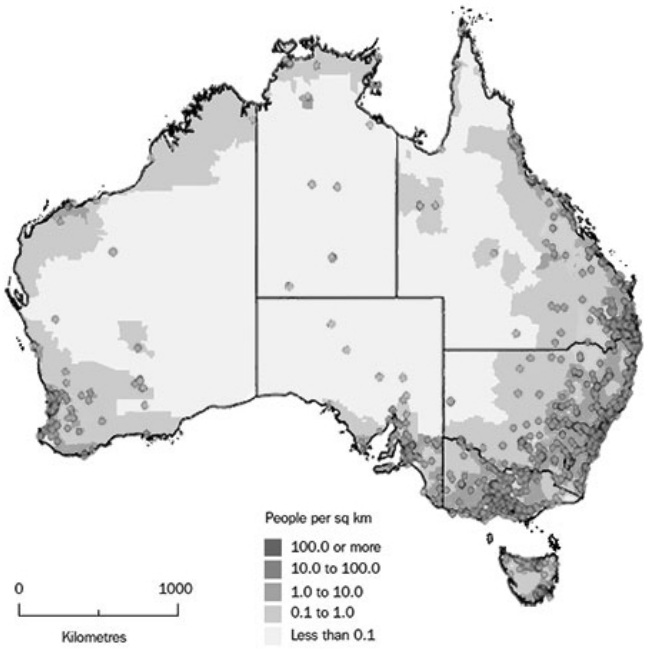
Sociodemographic Characteristics

Comparison with the 2011 Census data (Table 1) indicates that the study participants were broadly representative in their geographical distribution across Australia (both in terms of State or Territory and area of residence): three-quarters of young women (75.28%, 12,849/17,069) resided in major cities, compared to 74.5% among the population. Similarly to young women in the Census, the vast majority (95.62%, 16,321/17,069) had never been married. The age distribution of the study participants was also close to that of the population.

The main difference identified was that study participants had higher levels of educational attainment, for instance, only 7.45% (1271/17,069) had not completed year 12 (compared with 14.9% of the women in population in this age group). Slightly more than one-third of women (35.18%, 5931/16,857) were studying 16 or more hours a week (Table 2).

Figure 1 illustrates the broad geographical distribution of the cohort (with each dot representing at least one individual) and reflects the relatively high population density along the East and South coast and the sparse population scattered across the central and northwest areas of the continent.

Figure 1. Distribution of women aged 18-23 years recruited using the Internet across Australia (N=17,069). Each dot represents at least 1 woman. Figure adapted from the Regional Population Growth, Australia (catalogue 3218.0.).



XSL•FO
RenderX

Table 1. Comparison of sociodemographic characteristics of women aged 18-23 years, recruited using the Internet, with 2011 Australian Census data.

Characteristic	Study participants, 2012-2013 (N=17,069)		Census data, 2011 (N=844,636)
	n[a] (%)	95% CI	%
Age			
18	2599 (15.2)	14.7-15.8	16.0
19	2986 (17.5)	17.0-18.1	16.2
20	2924 (17.1)	16.6-17.7	16.8
21	2809 (16.5)	15.9-17.0	17.1
22	2879 (16.9)	16.3-17.4	16.9
23	2851 (16.7)	16.2-17.3	17.0
State/Territories			
New South Wales	4741 (27.9)	27.2-28.5	31.0
Victoria	4089 (24.0)	23.4-24.7	25.4
Queensland	3807 (22.4)	21.8-23.0	20.6
Western Australia	1883 (11.1)	10.6-11.5	10.6
South Australia	1301 (7.6)	7.2-8.0	7.3
Australian Capital Territory	560 (3.3)	3.0-3.6	2.0
Tasmania	494 (2.9)	2.7-3.2	2.1
Northern Territory	138 (0.8)	0.7-0.9	1.0
Area of residence			
Major city	12,849 (75.3)	74.6-75.9	74.5
Inner regional	2831 (16.6)	16.0-17.1	16.0
Outer regional	1151 (6.7)	6.4-7.1	7.2
Remote	131 (0.8)	0.6-0.9	1.1
Very remote	52 (0.3)	0.2-0.4	0.9
Migratory/no usual address/missing	55 (0.3)	0.2-0.4	0.3
Highest qualification			
Less than Year 12	1271 (7.4)	7.1-7.9	14.9
Year 12 or equivalent	7341 (43.0)	42.8-44.2	46.1
Trade/certificate / diploma	4428 (26.0)	25.6-26.9	21.7
University degree	3844 (22.5)	21.1-23.4	9.4
Missing/not stated/ inadequately described	185 (1.2)	1.0-1.3	7.8
Aboriginal or Torres Strait Islander origin [b]			
No	14,529 (97.4)	97.2-97.7	91.9
Yes	384 (2.6)	2.3-2.8	3.5
Missing	-	-	4.7
Marital status			
Never married	16,321 (95.6)	95.2-95.9	94.5
Married	510 (3.0)	2.7-3.2	4.9
Separated/divorced/widowed	53 (0.3)	0.2-0.4	0.6
Missing	185 (1.2)	1.10-1.3	-

[a]Numbers may not sum to total due to missing data.

[b]N=14,913 as this question was only asked in a later version of the survey.

XSL•FO
RenderX

Table 2. Comparison of education level and lifestyle characteristics of women aged 18-23 years, recruited using the Internet and social networking sites, with the 2011-2012 Australian National Health Survey.

		Study participants 2012-2013 (N=17,069)		Study participants (weighted for education level)	Australian National Health Survey 2011-2013
		n[a] (%)	95% CI	% (95% CI)	%
Education level					
	Less than Year 12	1271 (7.4)	7.1-7.9	15.4 (14.6-16.1)	17.9
	Year 12 or equivalent	7341 (43.0)	42.8-44.2	49.7 (48.9-50.5)	40.1
	Certificate/diploma	4428 (26.0)	25.6-26.9	23.4 (22.8-24.1)	32.9
	University degree	3844 (22.5)	21.1-23.4	10.4 (10.0-10.7)	8.7
	Missing/not stated /inadequ descr	185 (1.2)	1.0-1.3	1.1 (1.0-1.3)	0.4
Self-rated health (weighted)					
	Excellent	1097 (6.5)	6.1-6.9	5.8 (5.5-6.2)	17.7
	Very good	6081 (36.0)	35.3-36.7	33.8 (33.0-34.5)	38.1
	Good	6866 (40.6)	39.9-41.4	41.7 (40.9-42.5)	34.4
	Fair/poor	2859 (16.9)	16.3-17.5	18.8 (18.1-19.4)	10.1
Smoking status					
	Non-smoker	13,701 (81.1)	80.5-81.7	78.3 (77.6-79.0)	85.0
	Current smoker	3188 (18.9)	18.3-19.5	21.8 (21.0-22.4)	16.4
Body mass index (kg/m^2; weighted)					
	Underweight (<18.5)	1332 (7.9)	7.5-8.3	7.8 (7.4-8.3)	5.4
	Normal weight (18.5-25)	9923 (59.1)	58.3-59.8	57.4 (56.6-58.2)	58.2
	Overweight (25-30)	3231 (19.2)	18.7-19.9	19.5 (18.9-20.2)	14.8
	Obese (>30)	2298 (13.7)	13.2-14.2	15.3 (14.6-15.9)	18.8
Mean weight (kg)		67.4	67.1-67.6	67.8 (67.5-68.0)	66.4
Mean height (cm)		166.2	166.1-166.4	166.1 (166.0-166.2)	164.6
Mean BMI		24.4	24.3-24.5	24.6 (24.5-24.7)	24.7
Physical activity level					
	Inactive	1024 (6.0)	5.6-6.4	6.7 (6.3-7.2)	9.8
	Insufficiently active	5631 (33.1)	32.4-33.8	34.0 (33.2-34.8)	41.3
	Sufficiently active	10,359 (60.9)	60.2-61.6	59.3 (58.5-60.1)	48.3
Alcohol consumption (weighted)					
	Never drink	731 (4.5)	4.2-4.8	5.0 (4.6-5.3)	12.7
	Less than once a month	5629 (34.4)	33.7-35.2	35.5 (34.7-36.3)	21.4
	Less than once a week	5572 (34.1)	33.2-34.8	33.6 (32.8-34.4)	31.9
	At least once a week	4417 (27.0)	26.3-27.7	26.0 (25.2-26.7)	34.0
Paid employment					
	Yes	13,156 (77.9)	77.3-78.6	74.9 (74.2-75.7)	-
	No	3721 (22.1)	21.4-22.7	25.1 (24.3-25.8)	-
Studying (≥16 hrs per week)					
	Yes	5931 (35.2)	34.5-35.9	33.6 (32.8-34.3)	-
	No	10,926 (64.8)	64.1-65.5	66.4 (65.7-67.2)	-

[a]Numbers may not sum to total due to missing data.

XSL•FO
RenderX

Health-Related and Lifestyle Characteristics

A higher percentage of study participants rated their health in the online survey as fair or poor than women participating in face-to-face interviews in the AHS (18.77%, 3203/17,069 from weighted data vs 10.1% respectively) (Table 2). Study participants reported higher rates of smoking (21.78%, 3718/17,069 vs 16.4% were current smokers) and physical activity (58.97%, 10,150/17,211 were classified as sufficiently active vs 48.3% in the AHS) but lower levels of alcohol consumption than women in the AHS (59.58%, 9865/16,558 vs 65.9% drank alcohol at least once per month). Based on self-reported height and weight, 19.54% (3313/16,956) of the participants were classified as overweight (25≤ BMI< 30), 15.26% (2588/16,956) were obese (BMI≥30), compared with 14.8% and 18.8% respectively in the AHS.

The percentages obtained from unweighted data were similar to the education-weighted data and did not make substantive changes in the comparisons with AHS data.

Discussion

Principal Results

This study examines the representativeness, in terms of sociodemographic and lifestyle characteristics of a national cohort of young Australian women (born in 1989-95) who were recruited in 2012-13 mainly via social media and other Internet platforms, and completed the baseline survey online. The area of residence of the study participants is broadly representative of the geographical distribution of the population. The main sociodemographic difference was the higher proportion of women who had post-secondary school qualifications. The percentages for current smokers and those who were physically active among study participants were higher than the findings from the AHS. Based on self-reported height and weight, more than one-third of young women were identified as overweight or obese, similar to the percentage in these combined BMI categories found in the AHS. However, within these categories there were differences between studies: a higher percentage of study participants was overweight than in the AHS and a lower percentage was obese.

These comparisons used data from study participants weighted to match the education level of women in this age group in the national population. Little substantive difference in the distribution of health and lifestyle characteristics is evident when using unweighted data from study participants.

Limitations

There are a number of considerations to take into account when comparing prevalence estimates across population-based studies. For example, the eligibility criterion of the ALSWH limits study participants to those with the Medicare number, whereas the Census data are based the entire population including visitors, and the AHS survey was a sample of those living in Australia for at least 1 year or with the intention of doing so. Thus, unlike ALSWH, both the Census and AHS data included the sizable number of women students from overseas studying in Australia [43]. It is also possible that some of the women may have based their educational level on their current studies (for a university degree or trade qualification) rather than their completed qualifications. Over-representation of participants with higher educational levels is also recognized as a characteristic of many epidemiological studies [4,35], including the previously recruited cohorts in ALSWH. Also, AHS has a sample size of women in the similar age group of about 1000, considerably smaller than the number of study participants in this age group in ALSWH.

Another issue concerns differences in the mode of survey administration. For instance, the AHS was conducted via a face-to-face interview [43], whereas the study participants completed an online questionnaire. This difference may have a varying degree of impact according to the nature of the survey question and social or cultural factors that may influence an open response among some women, such as for levels of alcohol consumption. Furthermore, variations in the wording of questions or the available response options, such as for physical activity level, may limit the comparability of results. Yet it is worth noting that even though the BMI data for the study participants were calculated from self-assessed height and weight, whereas for the AHS height and weight were measured directly, the proportion of those classified as overweight or obese were similar in both studies. This is consistent with a previous study that found that Web-based self-reported data provide a valid measure of weight [44].

While 2.57% (384/14,913) of the study participants identified themselves as Aboriginal or Torres Strait Islander women, this is lower than the 3.5% from the 2011 Census data. This was expected, since effective recruitment and retention of participants from the Indigenous population requires culturally specific protocols that are best implemented in a separate and specially designed study.

Comparison With Prior Work

Previous small scale studies have used online social networking sites for recruitment but this is one of the few, population-based studies to rely on dynamically targeted advertising through Facebook to recruit a large cohort of young women for a national longitudinal health study. Other studies have also examined the cost-effectiveness of Facebook for recruitment [4,26,45], but only one study by Fenner et al [4] reported on the characteristics of the subsequent sample. Although for a smaller scale study, Fenner et al [4] also targeted young Australian women via Facebook, using separate advertising campaigns to target different age groups and regions, and were successful in obtaining a broadly representative of young women.

Implications

The representativeness of the sample in terms of key attributes, such as sociodemographic characteristics, is necessary to maximize external validity and strengthen the evidence base for policy and health care planning [46]. It is central to describing the health of a population at a particular point in time.

It is not the only consideration, however, as sufficient size and heterogeneity of the sample are important attributes for research on the relationships between risk factors and health outcomes and patterns over time that can provide insights on the

XSL•FO
RenderX

underlying causal mechanisms at work. Size and diversity of the sample are also important for identifying the health status and health care needs of minority groups.

Conclusions

Findings from this study support the use of the Internet and social networking sites as a viable recruitment method for large heterogeneous samples of young adults who are broadly representative of the population. Researchers need to be mindful that given the rapidly changing landscape of online social media, the exact strategies likely to be most effective for recruitment may also vary over time and according to the targeted subpopulation or age group of interest.

Acknowledgments

The research on which this paper is based was conducted as part of the Australian Longitudinal Study on Women's Health, the University of Newcastle and the University of Queensland. We are grateful to the Australian Government Department of Health for funding and to the women who provided the survey data. GDM is funded by the Australian Research Council Future Fellowship (FT120100812). IR was supported by the Australian National Health and Medical Research Council (grant number: APP1000986).

Authors' Contributions

GDM and DL contributed to design of the study and GDM drafted the manuscript. RH and JP conducted the statistical analyses. DL, LT, IR, and JEB contributed to the interpretation of the results and critical revision of the manuscript. AJD contributed to all stages of the study, interpretation of the results, and critical revision of the manuscript for important intellectual content. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Galea S, Tracy M. Participation rates in epidemiologic studies. Ann Epidemiol 2007 Sep;17(9):643-653. [doi: 10.1016/j.annepidem.2007.03.013] [Medline: 17553702]
2. van Gelder MM, Bretveld RW, Roeleveld N. Web-based questionnaires: the future in epidemiology? Am J Epidemiol 2010 Dec 1;172(11):1292-1298 [FREE Full text] [doi: 10.1093/aje/kwq291] [Medline: 20880962]
3. Kesse-Guyot E, Andreeva V, Castetbon K, Vernay M, Touvier M, Méjean C, et al. Participant profiles according to recruitment source in a large Web-based prospective study: experience from the Nutrinet-Santé study. J Med Internet Res 2013 Sep;15(9):e205 [FREE Full text] [doi: 10.2196/jmir.2488] [Medline: 24036068]
4. Fenner Y, Garland SM, Moore EE, Jayasinghe Y, Fletcher A, Tabrizi SN, et al. Web-based recruiting for health research using a social networking site: an exploratory study. J Med Internet Res 2012 Feb;14(1):e20 [FREE Full text] [doi: 10.2196/jmir.1978] [Medline: 22297093]
5. Harris ML, Herbert D, Loxton D, Dobson A, Wigginton B, Lucke J. Recruiting young women for health surveys: traditional random sampling methods are not cost-effective. Aust N Z J Public Health 2014 Oct;38(5):495. [doi: 10.1111/1753-6405.12281] [Medline: 25269982]
6. Alshaikh F, Ramzan F, Rawaf S, Majeed A. Social network sites as a mode to collect health data: a systematic review. J Med Internet Res 2014 Jul;16(7):e171 [FREE Full text] [doi: 10.2196/jmir.3050] [Medline: 25048247]
7. Rothman KJ, Gallacher JE, Hatch EE. Rebuttal: When it comes to scientific inference, sometimes a cigar is just a cigar. Int J Epidemiol 2013 Aug;42(4):1026-1028 [FREE Full text] [doi: 10.1093/ije/dyt124] [Medline: 24062292]
8. Ebrahim S, Davey Smith G. Commentary: Should we always deliberately be non-representative? Int J Epidemiol 2013 Aug;42(4):1022-1026 [FREE Full text] [doi: 10.1093/ije/dyt105] [Medline: 24062291]
9. Rothman KJ, Gallacher JE, Hatch EE. Why representativeness should be avoided. Int J Epidemiol 2013 Aug;42(4):1012-1014 [FREE Full text] [doi: 10.1093/ije/dys223] [Medline: 24062287]
10. Stang A, Jöckel KH. Avoidance of representativeness in presence of effect modification. Int J Epidemiol 2014 Apr;43(2):630-631. [doi: 10.1093/ije/dyt263] [Medline: 24408970]
11. Nohr EA, Olsen J. Commentary: Epidemiologists have debated representativeness for more than 40 years--has the time come to move on? Int J Epidemiol 2013 Aug;42(4):1016-1017 [FREE Full text] [doi: 10.1093/ije/dyt102] [Medline: 24062289]
12. Richiardi L, Pizzi C, Pearce N. Commentary: Representativeness is usually not necessary and often should be avoided. Int J Epidemiol 2013 Aug;42(4):1018-1022 [FREE Full text] [doi: 10.1093/ije/dyt103] [Medline: 24062290]
13. Elwood JM. Commentary: on representativeness. Int J Epidemiol 2013 Aug;42(4):1014-1015 [FREE Full text] [doi: 10.1093/ije/dyt101] [Medline: 24062288]
14. Schooling CM, Jones HE. Is representativeness the right question? Int J Epidemiol 2014 Apr;43(2):631-632 [FREE Full text] [doi: 10.1093/ije/dyt264] [Medline: 24436097]

XSL•FO
RenderX

15. Andreeva VA, Galan P, Julia C, Castetbon K, Kesse-Guyot E, Hercberg S. Assessment of response consistency and respective participant profiles in the Internet-based NutriNet-Santé Cohort. Am J Epidemiol 2014 Apr 1;179(7):910-916. [doi: 10.1093/aje/kwt431] [Medline: 24521560]
16. Bonn SE, Trolle Lagerros Y, Christensen SE, Möller E, Wright A, Sjölander A, et al. Active-Q: validation of the web-based physical activity questionnaire using doubly labeled water. J Med Internet Res 2012 Feb;14(1):e29 [FREE Full text] [doi: 10.2196/jmir.1974] [Medline: 22356755]
17. da Costa FF, Schmoelz CP, Davies VF, Di Pietro PF, Kupek E, de Assis MA. Assessment of diet and physical activity of Brazilian schoolchildren: usability testing of a web-based questionnaire. JMIR Res Protoc 2013 Aug;2(2):e31 [FREE Full text] [doi: 10.2196/resprot.2646] [Medline: 23958804]
18. Smith JA, Wilde MH, Brasch J. Internet recruitment and retention for a 6 months' longitudinal study. J Nurs Scholarsh 2012 Jun;44(2):165-170. [doi: 10.1111/j.1547-5069.2012.01446.x] [Medline: 22551052]
19. Hercberg S, Castetbon K, Czernichow S, Malon A, Mejean C, Kesse E, et al. The Nutrinet-Santé Study: a web-based prospective study on the relationship between nutrition and health and determinants of dietary patterns and nutritional status. BMC Public Health 2010;10:242 [FREE Full text] [doi: 10.1186/1471-2458-10-242] [Medline: 20459807]
20. Russell CW, Boggs DA, Palmer JR, Rosenberg L. Use of a web-based questionnaire in the Black Women's Health Study. Am J Epidemiol 2010 Dec 1;172(11):1286-1291 [FREE Full text] [doi: 10.1093/aje/kwq310] [Medline: 20937635]
21. Australian Longitudinal Study on Women's Health (ALSWH). URL: http://alswh.org.au/for-participants/1989-95-cohort [accessed 2014-07-21] [WebCite Cache ID 6RDkFTp6J]
22. Pew Research Center. The Web at 25. URL: http://www.pewinternet.org/2014/02/27/the-web-at-25-in-the-u-s/ [accessed 2014-08-14] [WebCite Cache ID 6RoWcDAIv]
23. Zickuhr K, Smith A. Pew Research Center 2012. Washington, DC: Pew Internet and American Life Project Digital differences URL: http://www.pewinternet.org/files/old-media//Files/Reports/2012/PIP_Digital_differences_041312.pdf [accessed 2014-08-14] [WebCite Cache ID 6RoXEft8y]
24. Duggan M, Smith A. Pew Research Center. 2014. Social Media Update 2013 URL: http://www.pewinternet.org/2013/12/30/social-media-update-2013/ [accessed 2014-08-14] [WebCite Cache ID 6RoVRkYCK]
25. Godfrey M. Facebook checked by 9 million Australians every day. Sydney, Australia: The Sydney Morning Herald; 2013 Aug 20. URL: http://www.smh.com.au/digital-life/digital-life-news/facebook-checked-by-9-million-australians-every-day-20130820-2s7wo.html [accessed 2014-11-16] [WebCite Cache ID 6U7VKBHTM]
26. Ramo DE, Prochaska JJ. Broad reach and targeted recruitment using Facebook for an online survey of young adult substance use. J Med Internet Res 2012 Feb;14(1):e28 [FREE Full text] [doi: 10.2196/jmir.1878] [Medline: 22360969]
27. 8146.0 - Household Use of Information Technology, Australia, 2012-13. Canberra, Australia: Australian Bureau of Statistics (ABS) URL: http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/8A12E6E0D07D36A0CA257C89000E3FB7?opendocument [accessed 2014-07-21] [WebCite Cache ID 6RDkahu13]
28. The social and psychological impact of online social networking.: Australian Psychological Society (APS); 2010. URL: http://www.psychology.org.au/publications/inpsych/2010/december/social/ [accessed 2014-07-21] [WebCite Cache ID 6RDkoqNfG]
29. Berkman LF, Glass T, Brissette I, Seeman TE. From social integration to health: Durkheim in the new millennium. Soc Sci Med 2000 Sep;51(6):843-857. [Medline: 10972429]
30. Richiardi L, Pivetta E, Merletti F. Recruiting study participants through Facebook. Epidemiology 2012 Jan;23(1):175. [doi: 10.1097/EDE.0b013e31823b5ee4] [Medline: 22157313]
31. Lord S, Brevard J, Budman S. Connecting to young adults: an online social network survey of beliefs and attitudes associated with prescription opioid misuse among college students. Subst Use Misuse 2011 Jan;46(1):66-76 [FREE Full text] [doi: 10.3109/10826084.2011.521371] [Medline: 21190407]
32. Lohse B. Facebook is an effective strategy to recruit low-income women to online nutrition education. J Nutr Educ Behav 2013 Jan;45(1):69-76. [doi: 10.1016/j.jneb.2012.06.006] [Medline: 23305805]
33. Kapp JM, Peters C, Oliver DP. Research recruitment using Facebook advertising: big potential, big challenges. J Cancer Educ 2013 Mar;28(1):134-137. [doi: 10.1007/s13187-012-0443-z] [Medline: 23292877]
34. Amon KL, Campbell AJ, Hawke C, Steinbeck K. Facebook as a recruitment tool for adolescent health research: a systematic review. Acad Pediatr 2014 Sep;14(5):439-447.e4. [doi: 10.1016/j.acap.2014.05.049] [Medline: 25169155]
35. Lee C, Dobson AJ, Brown WJ, Bryson L, Byles J, Warner-Smith P, et al. Cohort profile: the Australian Longitudinal Study on Women's Health. Int J Epidemiol 2005 Oct;34(5):987-991 [FREE Full text] [doi: 10.1093/ije/dyi098] [Medline: 15894591]
36. Brown WJ, Dobson AJ, Bryson L, Byles JE. Women's Health Australia: on the progress of the main cohort studies. J Womens Health Gend Based Med 1999 Jun;8(5):681-688. [Medline: 10839654]
37. 2011.0.55.001 - Information Paper: Census of Population and Housing -- Products and Services Canberra, Australia. Canberra, Australia: Australian Bureau of Statistics (ABS); 2011. URL: http://www.abs.gov.au/ausstats/abs@.nsf/lookup/2011.0.55.001Main%20Features22011 [accessed 2014-08-14] [WebCite Cache ID 6RoXZ1CrZ]

XSL•FO
RenderX

38. 4364.0.55.003 - Australian Health Survey: Updated Results, 2011-2012. Canberra, Australia: Australian Bureau of Statistics (ABS) URL: http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/4364.0.55.003Chapter1002011-2012 [accessed 2014-08-14] [WebCite Cache ID 6RoXqZ20i]
39. Australian Institute of Health and Welfare (AIHW). Rural, regional and remote health: a guide to remoteness classifications. Canberra: AIHW; 2004.
40. World Health Organization (WHO). Obesity: preventing and managing the global epidemic. Geneva: WHO; 2000.
41. National Health and Medical Research Council. Australian guidelines to reduce health risks from drinking alcohol. Canberra: Commonwealth of Australia; 2009.
42. Australian Institute of Health and Welfare (AIHW). The Active Australia Survey: A guide and manual for implementation, analysis and reporting. Canberra: AIHW; 2003. URL: http://www.aihw.gov.au/publication-detail/?id=6442467449 [accessed 2014-07-21] [WebCite Cache ID 6RDnB7FuL]
43. Australian Bureau of Statistics (ABS). 4363.0.55.001 - Australian Health Survey: Users' Guide, 2011-13. Canberra: ABS; 2013. URL: http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/D1A83EE0703D6E8FCA257BBB0014C443?opendocument [accessed 2014-07-21] [WebCite Cache ID 6RDnLfqcI]
44. Bonn SE, Trolle Lagerros Y, Bälter K. How valid are Web-based self-reports of weight? J Med Internet Res 2013 Apr;15(4):e52 [FREE Full text] [doi: 10.2196/jmir.2393] [Medline: 23570956]
45. Lohse B, Wamboldt P. Purposive facebook recruitment endows cost-effective nutrition education program evaluation. JMIR Res Protoc 2013 Aug;2(2):e27 [FREE Full text] [doi: 10.2196/resprot.2713] [Medline: 23948573]
46. Glasgow RE, Emmons KM. How can we increase translation of research into practice? Types of evidence needed. Annu Rev Public Health 2007 Apr;28:413-433. [doi: 10.1146/annurev.publhealth.28.021406.144145] [Medline: 17150029]

Abbreviations

AHS: Australian Health Survey
ALSWH: Australian Longitudinal Study on Women's Health
BMI: body mass index

Edited by G Eysenbach; submitted 17.08.14; peer-reviewed by J Wark, E Hatch; comments to author 11.09.14; revised version received 26.09.14; accepted 17.10.14; published 15.12.14.

Please cite as:
Mishra GD, Hockey R, Powers J, Loxton D, Tooth L, Rowlands I, Byles J, Dobson A
Recruitment via the Internet and Social Networking Sites: The 1989-1995 Cohort of the Australian Longitudinal Study on Women's Health
J Med Internet Res 2014;16(12):e279
URL: http://www.jmir.org/2014/12/e279/
doi:10.2196/jmir.3788
PMID:25514159

©Gita Devi Mishra, Richard Hockey, Jennifer Powers, Deborah Loxton, Leigh Tooth, Ingrid Rowlands, Julie Byles, Annette Dobson. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 15.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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 an Electronic Patient Portal Among the Chronically Ill: An Observational Study

Iiris Riippa[1], MSc (Econ), MSc (Tech); Miika Linna[1], DSc (Tech); Ilona Rönkkö[2], RN, M Health Care; Virpi Kröger[2], MD

[1]Aalto University, Espoo, Finland

[2]Hämeenlinnan Terveyspalvelut Public Utility, Hämeenlinna, Finland

Corresponding Author:
Iiris Riippa, MSc (Econ), MSc (Tech)
Aalto University
Otaniementie 17
Espoo,
Finland
Phone: 358 503430063
Fax: 358 947023665
Email: iiris.riippa@aalto.fi

Abstract

Background: Electronic patient portals may enhance effective interaction between the patient and the health care provider. To grasp the full potential of patient portals, health care providers need more knowledge on which patient groups prefer electronic services and how patients should be served through this channel.

Objective: The objective of this study was to assess how chronically ill patients' state of health, comorbidities, and previous care are associated with their adoption and use of a patient portal.

Methods: A total of 222 chronically ill patients, who were offered access to a patient portal with their health records and secure messaging with care professionals, were included in the study. Differences in the characteristics of non-users, viewers, and interactive users of the patient portal were analyzed before access to the portal. Patients' age, gender, diagnoses, levels of the relevant physiological measurements, health care contacts, and received physiological measurements were collected from the care provider's electronic health record. In addition, patient-reported health and patient activation were assessed by a survey.

Results: Despite the broad range of measures used to indicate the patients' state of health, the portal user groups differed only in their recorded diagnosis for hypertension, which was most common in the non-user group. However, there were significant differences in the amount of care received during the year before access to the portal. The non-user group had more nurse visits and more measurements of relevant physiological outcomes than viewers and interactive users. They also had fewer referrals to specialized care during the year before access to the portal than the two other groups. The viewers and the interactive users differed from each other significantly in the number of nurse calls received, the interactive users having more calls than the viewers. No significant differences in age, gender, or patient activation were detected between the user groups.

Conclusions: Previous care received by the patient is an important predictor for the use of a patient portal. In a group of patients with a similar disease burden, demand for different types of health services and preferences related to the service channel seem to contribute to the choice to use the patient portal. Further research on patient portal functionalities and their potential to meet patient needs by complementing or substituting for traditional health care services is suggested.

(J Med Internet Res 2014;16(12):e275) doi:10.2196/jmir.3722

KEYWORDS

chronic illness; patient portal; service channel; user profile

Introduction

The electronic patient portal is an increasingly popular channel for health care providers to offer information to and interact with their patients. Typically, a patient portal includes patients' own health records, drawn from the care provider's electronic health records (EHR), and the possibility of interacting with the provider through secure messaging in non-acute matters and to

XSL•FO
RenderX

request repeat prescriptions [1]. More advanced portals may also offer personally tailored health information and social functionalities that enable peer support from other patients [2].

The potential benefits of patient portals include the empowerment and activation of patients in the management of their own health through increased access to related information [2,3]. In addition, interaction through a patient portal may improve the efficiency of care by replacing some of the service contacts previously performed in person or via phone calls, which are thus bound to time and often to place [4].

The suggested benefits of the patient portals may, however, be unequally distributed among patients, owing to differing interest in, access to, or ability to use the service [5,6]. To grasp the full potential of patient portals and to avoid unequal access to care, health care providers need to know which of their patients may be served by this means, and how. This understanding may be acquired by identifying differences in characteristics between portal adopters and non-adopters and between users and non-users of specific functionalities. Previous studies have reported disparities in patient use of health information technology mostly by sociodemographic factors [7]. As the reported associations between sociodemographic factors and use of health information technology have been contradictory [7], other patient factors may explain the differential use of specific services offered by means of health information technology. In studies on patient portal use, little attention has been paid to patients' prior health care consumption patterns, that is, care received by the patients prior to portal access.

This study focused on disparities in patient portal use by patients' state of health and previous care received. As the adoption of a patient portal requires some level of patient participation, we also tested for the relationship between patient activation (knowledge, skills, and confidence in managing one's condition) and portal use. In addition, the associations of age and gender with portal use were assessed.

In chronic care, "frequent interactions with the provider [are] required and sustained effort [is] needed of the patient to manage his/her disease" [8]. It is therefore suggested that the chronically ill are likely to benefit from the use of an electronic patient portal [8]. Consequently, the present study was conducted among the chronically ill.

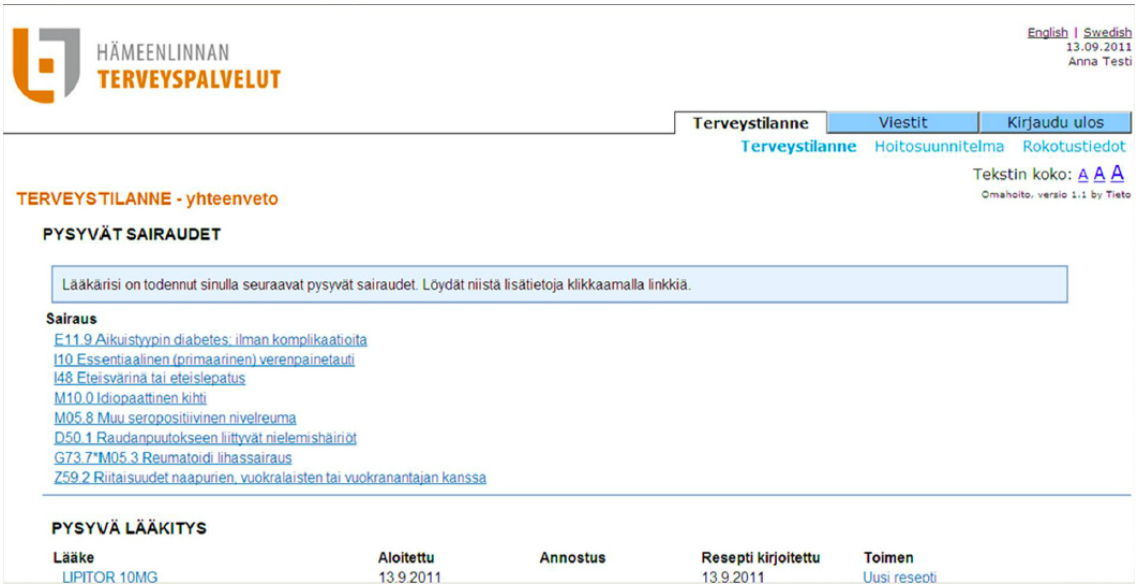
Methods

Study Setting, Participants, and Description of the Portal

The study setting was Finnish public primary care in a medium-sized Finnish town, Hämeenlinna, with c. 68,000 inhabitants and 10 health centers. In Finland, health services are mainly funded by municipalities from tax revenue. Municipalities are responsible for providing all necessary health services to their residents. Typically, primary care services are provided locally in health centers, whereas hospital districts formed by municipalities are responsible for arranging specialized medical care that is centralized in larger towns. Follow-up and maintenance of the chronically ill are one of the main tasks of public primary care in Finland.

The study group consisted of chronically ill, existing customers of the care provider, who had participated in a controlled before/after study reported elsewhere [9]. The eligibility criteria for the participants were: (1) age of at least 18 years, (2) has at least two treatable health conditions assessed by a health professional, (3) has bank identifiers for electronic identification and access to the Internet, and (4) is willing and able, both according to themselves and to a health care professional, to engage in using the portal.

Figure 1. Screenshot of the patient portal.



The eligible patients were offered access to the patient portal during their visit to primary health care facilities (Figure 1). During the visit, they were given information on the contents and use of the portal. Patients could access the portal through the health care provider's home page, using their bank identifiers for electronic identification. The functionalities of the portal included viewing the patient's personal health record containing diagnoses, medication prescriptions, and laboratory results,

XSL•FO
RenderX

viewing a personal care plan that the patient and a nurse had drawn up together during an in-person visit in order to holistically care for the patient's health, electronic messaging with the care team, and prescription renewal.

Informed consent to participate was collected from each participant. Ethical approval was granted by the ethical board of the local authority (Pirkanmaa Hospital District).

Materials

Owing to the diagnostic heterogeneity of the study group, three different types of diagnosis indicators were gathered to represent the comorbidity of the patients. First, diagnoses for the most common chronic illnesses in the study group, namely type 1 or 2 diabetes, hypertension, and hypercholesterolemia, were identified through the International Classification of Diseases (ICD-10) codes. Second, the Charlson Comorbidity Index (CCI) was used to assess the comorbidity of the patients. The CCI is a widely used system for characterizing patient comorbidities, drawing on ICD-10 recordings of 17 common chronic medical conditions (Multimedia Appendix 1) [10]. Third, the total number of diseases identified as chronic by a health care professional were collected from patients' EHR.

The indicators for patients' state of health were physiological health outcomes and health as reported by the patients. Values of the physiological measures glycated hemoglobin (HbA1c), low-density lipoprotein (LDL), body mass index (BMI), and blood pressure were collected to assess the patients' state of physiological health. These measures were chosen because of their relevance in the management of the above-mentioned most common diseases in the participant group. The patients' reported health was assessed through the short-form health survey SF-36, a broadly used instrument that generates functionality-based scores for mental and physical health and wellness [11-13].

The care received by the patients was obtained from the EHR databases. Using the unique personal ID code, the service contacts and the physiological outcomes measurement in the year preceding access to the portal were retrieved. Service contacts included doctor and nurse visits and calls in primary care and referrals to specialized care. To assess the monitoring, we collected the measurements of the patients' relevant physiological health outcomes (HbA1c, BMI, LDL, and blood pressure).

To measure patient activation, we used the short form of Patient Activation Measure (PAM13), created by Judith Hibbard and colleagues [14], a validated instrument that assesses patient knowledge, skills, and confidence for self-management [14].

With the exception of SF-36 and PAM13 scores, all indicators were collected from the health care providers' EHR. An email with a link to a survey was sent to participants to collect their responses to SF-36 and PAM surveys at the time of offering access to the portal, and 150 participants responded. The patient portal's log information was collected to assess the use of the portal's functionalities.

Statistical Analysis

To analyze the predictors of patient portal use, participants were divided into three groups (non-users, viewers, and interactive users), based on their portal use during the 6 months after gaining access to the portal. Non-users did not log in to the portal during the follow-up period. Viewers logged in at least once, but did not use either of the interactional functionalities, namely messaging with the care team or prescription renewal. Interactive users logged in to the portal and used one of the interactional functionalities at least once.

For categorical variables, chi-square tests for overall differences among the three groups were used. To further identify such differences, pairwise comparisons using chi-square tests were conducted. Owing to the non-normality of the distributions for continuous variables, the non-parametric Kruskal-Wallis test was used to analyze overall differences among the three groups and the Wilcoxon-Mann-Whitney test for pairwise comparisons.

All statistical analyses were performed using Stata version 13 (StataCorp LP, College Station, TX). We used a CHARLSON Stata module by Stagg [15] to identify the CCI conditions from patient records and define the index value for each patient.

Results

Participants

A total of 876 patients visiting the health center facilities during the recruitment period from October 2011 to March 2012 fulfilled the eligibility criteria and were asked to participate in the study. Of these, 222 patients (25.3%) returned their informed consent and were included in the study. The mean age of patients was 62.7 (SD 9.0) years, and 49.1% (109/222) of them were women. The most frequent diagnoses of the study participants were type 1 or 2 diabetes, hypertension, and hypercholesterolemia. The majority of the patients had one disease or no diseases included in the CCI. The participants had visited a primary care doctor 3.4 (SD 3.2) times and a nurse 4.1 (SD 7.3) times, on average during the year before access to the portal (Table 1).

Use of Patient Portal Functionalities

Once they logged in to the portal, patients would encounter the starting page, containing their own health information, including diagnoses, medication prescriptions, and laboratory results. On average, this information was viewed 17.0 (SD 20) times per patient during the first year after access to the portal. The second most popular viewed feature of the portal, used 4.5 (SD 6.0) times on average, was their own personal care plan. On average, patients sent 2.1 (SD 3.5) messages to their care team and viewed their vaccination record 1.6 (SD 1.9) times. Only 0.4 (SD 1.1) prescription renewals per patient were made through the portal during the first year after access (Table 2).

XSL•FO
RenderX

Table 1. Descriptive characteristics of the study participants.

Characteristic		Total (n=222)
Age, mean (SD)		62.7 (9.0)
Female		109 (49.1%)
Most frequent diagnoses		
	Type 1 or 2 diabetes[ab]	103 (46.4%)
	Hypertension[ac]	96 (43.2%)
	Hypercholesterolemia[ad]	139 (62.6%)
Charlson comorbidity index [a]		
	0	93 (41.9%)
	1	79 (35.6%)
	2	50 (22.5%)
Office visits [e] **, mean (SD)**		
	Doctor visits	3.4 (3.2)
	Nurse visits	4.1 (7.3)

[a]From the time before access to the portal.

[b]ICD10 codes E10-E14 or ICPC codes T89-T90.

[c]ICD10 codes I10-I15 or ICPC codes K85-K87.

[d]ICD10 codes E78 or ICPC T93.

[e]During the year before access to the portal.

Table 2. Use of patient portal functionalities during the year after patient access.

Functionality	Mean (n=222)	SD	Min.	Max.
Viewing personal health record	17.0	20.0	0	146
Viewing personal care plan	4.5	6.0	0	44
Messages to the care team	2.1	3.5	0	25
Viewing vaccination record	1.6	1.9	0	13
Prescription renewal	0.4	1.1	0	7

Characteristics Predicting Patient Portal Use

The overall differences in age, gender, diagnoses, health outcomes, received monitoring, service contacts, and patient activation among the three groups are presented in Table 3. Multiple pairwise comparisons for the same characteristics are presented in Multimedia Appendix 2.

The most significant differences between non-users, viewers, and interactive users were detected in monitoring and service contacts. The proportion of patients who had had their HbA1c measured and recorded during the year before access to the portal differed significantly between the groups (chi-square test, *P*=.03), being higher in the non-user group than among the viewers (pairwise chi-square test, *P*=.02) and the interactive user group (pairwise chi-square test, *P*=.01). Compared to the viewer group, the non-users were also more likely to have had their BMI (pairwise chi-square test, *P*=.02) and blood pressure (pairwise chi-square test, *P*=.02) measured and recorded. The non-users had visited a nurse most often (Kruskal-Wallis test, *P*=.01) but had fewer referrals to specialized care than the viewers (Wilcoxon-Mann-Whitney test, *P*=.02) and the interactive users (Wilcoxon-Mann-Whitney test, *P*=.03). The viewers and the interactive users differed from each other in the number of nurse calls. The interactive users had received more nurse calls during the preceding year than the viewers (Wilcoxon-Mann-Whitney test, *P*=.03).

The only comorbidity indicator that differed between the groups was the prevalence of hypertension diagnosis. The non-users were most likely to have a hypertension diagnosis (Kruskal-Wallis test, *P*=.01). There were no significant differences in mean age, gender distribution, or patient activation among the user profile groups. The statistically significant differences (*P*<.05) in patient characteristics among the user groups are presented in Table 4.

XSL•FO
RenderX

Table 3. Overall differences in patient characteristics among non-users, viewers, and interactive users.

Characteristic		Non-users (n=37)	Viewers (n=91)	Interactive users (n=94)	*P* value for difference among groups[h]
Female, n (%)		20 (54.1%)	44 (48.4%)	45 (47.9%)	.80
Age, years, mean (median, SD)		62.0 (64, 11.6)	63.0 (65, 8.8)	62.8 (65, 8.1)	.88
Comorbidity, n (%)					
	Type 1 or 2 diabetes[ab]	20 (54.1%)	36 (39.6%)	47 (50.0%)	.22
	Hypertension[ac]	24 (64.8%)	35 (38.5%)	37 (39.4%)	.01
	Hypercholesterolemia[ad]	22 (59.5%)	62 (68.1%)	55 (58.5%)	.37
Charlson index, n (%)					.51
	0	15 (40.5%)	40 (44.0%)	38 (40.4%)	
	1	10 (27.0%)	34 (37.4%)	35 (37.2%)	
	2	12 (32.4%)	17 (18.7%)	21 (22.3%)	
Chronic diagnoses, mean (median, SD)		2.3 (2, 1.5)	1.9 (2, 1.8)	1.9 (2, 1.8)	.22
Physiological health outcomes, mean (median, SD)					
	Glycated hemoglobin (HbA1c)[g]	42.4 (41.25, 9.4)	41.9 (39.8, 8.4)	44.2 (40.3, 12.7)	.96
	Low-density lipoprotein (LDL)[g]	2.7 (0.8)	3.0 (1.0)	3.0 (1.0)	.25
	Body mass index (BMI)[g]	31.0 (32.0, 5.6)	31.7 (31.0, 6.7)	31.5 (30.0, 6.8)	.98
	Blood pressure, diastolic[g]	82.2 (84.5, 11.8)	86.5 (87.5, 10.5)	86.4 (86.5, 10.6)	.17
	Blood pressure, systolic[g]	139.7 (140.3, 16.3)	143.7 (142.0, 21.4)	144.1 (145.0, 17.7)	.50
Patient-reported health, mean (median, SD)					
	SF-36 Physical Health at access	63.2 (66.0, 22.1)	65.9 (69.0, 20.4)	63.5 (66.5, 20.6)	.77
	SF-36 Mental Health at access	75.9 (80.0, 16.3)	75.8 (80.0, 19.0)	71.0 (78.5, 22.6)	.59
Monitoring, n (%)					
	At least one HbA1c measurement[e]	31 (83.8%)	57 (62.6%)	57 (60.6%)	.03
	At least one LDL measurement[e]	32 (2.48, 86.5%)	74 (2.83, 81.3%)	77 (2.81, 81.9%)	.77
	At least one BMI measurement[e]	31 (83.8%)	57 (62.6%)	66 (70.2%)	.06
	At least one blood pressure measurement[e]	34 (91.9%)	67 (73.63%)	77 (81.9%)	.05
Service contacts, mean (median, SD)					
	Doctor visits [e]	3.4 (2, 3.5)	2.9 (2, 2.6)	3.9 (3, 3.7)	.27
	Nurse visits[e]	5.2 (4, 3.9)	3.7 (3, 3.1)	4.7 (3, 10.8)	.01
	Doctor calls[e]	1.1 (0, 1.6)	1.3 (1, 1.6)	1.5 (1, 1.8)	.50
	Nurse calls[e]	1.0 (0, 1.3)	0.8 (0, 1.4)	1.0 (1, 1.7)	.13
	Referrals to secondary care[e]	0.2 (0, 0.7)	0.4 (0, 0.6)	0.5 (0, 0.8)	.15
Patient activation (PAM), mean (median, SD)					
	PAM score at access	63.5 (63.2, 11.7)	63.8 (66.0, 15.5)	62.4 (63.2, 15.1)	.84

[a]From the time before access to patient portal.

[b]ICD10 codes E10-E14 or ICPC codes T89-T90.

XSL•FO
RenderX

[c]ICD10 codes I10-I15 or ICPC codes K85-K87.

[d]ICD10 codes E78 or ICPC T93.

[e]During the year before access to patient portal.

[g]At least one measurement during the year before access to patient portal. If a patient had several measurements, the average is reported.

[h]Chi-square test for the categorical variables and non-parametric Kruskal-Wallis test for the continuous variables.

Table 4. Differences in patient characteristics between user groups.

Characteristics	Non-users	Viewers	Interactive users
Comorbidity	Most likely to have a hypertension diagnosis		
Monitoring	Most likely to have had their HbA1c measured and recorded		
	More likely than viewers to have had their BMI and BP measured and recorded	Less likely than non-users to have had their BMI and BP measured and recorded	
Service contacts	Most nurse visits		
	Least referrals to specialized care	Fewer nurse calls than interactive users	More nurse calls than viewers

Discussion

Principal Findings

In this study, we analyzed how patients' state of health, previous care received, age, gender, and patient activation predict the use of an electronic patient portal. The differences in these indicators were assessed between non-users, viewers, and interactive users. The main differences were detected in the previous care received by non-users and the two user groups that logged in to the portal.

Previous research has found positive [16-18] and negative [19,20] associations between patient portal use and use of other health care services. In this study, this association was found to vary by the type of health services previously received. Whereas the non-user group had visited a nurse most often, they had fewer referrals to specialized care than the two groups that had logged in to the portal. Among the patients who had logged in, the interactive users differed from the viewers in having received more nurse calls during the year before access. Plausible explanations for these observations may be found by considering the patient needs that an electronic patient portal potentially meets. In Finland, routine monitoring of the chronically ill is mostly performed by nurses. A lack of referrals to specialized care, combined with a higher number of nurse visits, may indicate a stable medical condition where patient needs are met and new channels for medical services are not needed. The association between higher numbers of nurse calls and interactive use of a patient portal may be explained by the nature of the interaction performed through these service channels. Compared to face-to-face visits with a health care provider, service encounters conducted by phone may be more apt for substitution by online interaction. Whereas these possible explanations are just some of many alternatives, the findings of this study do encourage, in line with Varsi and colleagues [21], a more fine-grained distinction between different types of health service encounters and respective patient-provider communication channels.

Unlike previous service use, state of health, age, gender, and patient activation had no significant association with patient portal use in this study. Previous findings on these associations are somewhat contradictory. In a study by Weingart and colleagues [19], patients who enrolled in a patient portal had fewer medical problems, and Tenforde and colleagues [22] found, among a diabetic cohort, that the users of a patient portal demonstrated better glycemic control. By contrast, the study by Earnest and colleagues [16] showed that the portal users were more symptomatic than the non-users. On the association between portal use and patient activation, Hibbard and Greene [23] stated that more activated patients are more likely to be referred to the patient portal and, among that group, the higher activated were more likely to actually use it, whereas Roblin and colleagues [24] found no significant association. In previous studies, younger [19,22] or older [18], and men [25] or women [24] have been suggested to be more likely or frequent users of patient portals.

The heterogeneity of the research results is likely due to the different settings of the studies. In particular, the chosen cohort and the functionalities offered through the portal may yield differing results. In this study, the participant group consisted of chronically ill patients, who were thus likely to benefit from the portal and were all explicitly offered access to it. Further, the portal itself included little functionality or information that might not be accessed at all via the traditional service channels, namely in-person visits and phone calls. Thus, it is not surprising that age, gender, patient activation, and state of health lost their relevance, whereas patients' demand for care and preferences in terms of service channel came to matter in the use of the patient portal. It should be noted, however, that only a minority of the study participants never logged in to the portal. This finding supports the suggestion that chronically ill patients are likely to benefit from, and thus to use, an electronic patient portal.

Some of the findings on predictors of patient portal use may also have reflected the contemporary novelty of health services

XSL•FO
RenderX

delivered through the Internet. Lately, it has been suggested that the digital divide between different sociodemographic groups due to lack of access to the Internet is narrowing [26]. In Finland in 2013, 92% of people aged 16 to 74 years and 85% of people aged 16 to 89 years had used the Internet in the past 3 months. Further, 79% of people aged 16 to 89 years had used Internet banking in the past 3 months [27]. It is thus unlikely that access to the Internet or the novelty of running errands online would, in general, hamper access to electronic patient portals, although health care-related online services are still something of a novelty in Finland. Rather, the results suggest that the non-user group did not perceive an additional benefit in using the portal, as they were already well served or they preferred the traditional service channels to electronic services. An important aspect of patient portal as a service channel is that, through it, receiving and providing information is not tied to time and place. To assess the value of this aspect to the patient, future research may benefit from applying behavioral research [28] or economic models on the individual demand for health services or health information in general [29].

In addition to the patient cohort targeted and the patient needs that are met through the portal, future research should pay attention to the influence of the personnel marketing the portal and its functionalities to the patients. In this study, the portal was explicitly offered to each participant, so the choice of use was left to the patient. Nevertheless, it is possible that, for example, some of the health centers involved in the study were encouraging use of the portal more than others. Further, while the prescription renewal functionality was offered in the portal, its use was not promoted, owing to the lack of a national prescriber-pharmacist interface at the time of the study.

Strengths and Limitations

This study contributes to previous research by extracting predictors for patients' choice to use an electronic patient portal in a group of patients that is likely to benefit from such portal, namely the chronically ill. Unlike in most studies with large patient cohorts [19], the portal was explicitly offered to each patient. The views of staff in the health care service on who may benefit from the portal did not therefore affect subsequent categorization of patients among user and non-user groups. Further, when analyzing the association between portal use and other services received, different contact types were specified. This provided a more refined view, typically neglected by previous studies [20]. Patients' state of health, which has an apparent connection with patients' need of and demand [30] for health care services, was assessed by several objective and subjective measures. This supports the validity of the conclusion that the groups did not differ in terms of disease burden but rather in their demand for different services or preference for different service channels.

Despite these contributions, the study also has limitations. A major and common limitation is the restriction of the study results to the patients who responded to the study request and who may fundamentally differ from those who decided not to participate. Further, the empiria of this study does not provide evidence of *why* the use of certain services is related to use of a patient portal. Neither was the effect of patient portal use on relevant care outcomes addressed. Further studies, analyzing different customer relationships and the relevance of a patient portal in these relationships, are needed to gain understanding on which functionalities of a patient portal may complement or substitute for traditional channels for service delivery in health care, and to assess the effect that this substitution may have on care outcomes.

Conclusions

In this study, the predictors of the use of an electronic patient portal were assessed among a group of patients likely to benefit from such a portal, namely the chronically ill. Previous care received by the patient, rather than state of health, age, gender, and patient activation, was an important factor predicting the attractiveness of electronic patient portal use.

Previous research on patient characteristics predicting the use of electronic patient portals has shown contradictory results. This is partially due to the differences in patient cohorts and portal functionalities. However, some of the predictors may also be losing their relevance as the novelty of online health care services levels out. As sociodemographic factors become less accurate predictors of online service use in Western countries, individual preferences in terms of service channel, as well as the functionalities offered through a patient portal, become relevant when identifying the potential uses of such a portal. To grasp the full potential of electronic patient portals, care providers need to know what types of services may be provided through a patient portal. Further research on patient portal functionalities and their potential to meet patient needs by complementing or replacing traditional health care services is suggested.

Acknowledgments

The authors wish to thank Tieto Corporation for collection of the data, and Dr Tuomo Lehtovuori and Dr Osmo Saarelma for their valuable comments on the study setting and on the practical implications of the results. The following professionals working at Hämeenlinnan Terveyspalvelut Public Utility deserve special recognition for their aid in realizing the research setting and data collection: CEO Risto Mäkinen (MD), Nursing Director Kirsti Helkiö (RN), Service Designers Jari Numminen (RN, BEng) and Sami Kiviharju (MSc), Mia Haapanen (RN), and Elina Pohja (RN). Finally, the authors are obliged to the nurses and doctors working in Hämeenlinnan Terveyspalvelut, who participated in the recruitment of the patients and made this study possible.

The study was partly supported by the SalWe Research Program for IMO (Tekes – the Finnish Funding Agency for Technology and Innovation grant 648/10).

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1

Charlson Comorbidity Index conditions.

[PDF File (Adobe PDF File), 27KB - jmir_v16i12e275_app1.pdf]

Multimedia Appendix 2

Pairwise comparisons of patient characteristics.

[PDF File (Adobe PDF File), 35KB - jmir_v16i12e275_app2.pdf]

References

1. 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]
2. 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]
3. Ross SE, Lin CT. The effects of promoting patient access to medical records: a review. J Am Med Inform Assoc 2003;10(2):129-138 [FREE Full text] [Medline: 12595402]
4. Zhou YY, Kanter MH, Wang JJ, Garrido T. Improved quality at Kaiser Permanente through e-mail between physicians and patients. Health Aff (Millwood) 2010 Jul;29(7):1370-1375 [FREE Full text] [doi: 10.1377/hlthaff.2010.0048] [Medline: 20606190]
5. Goel MS, Brown TL, Williams A, Hasnain-Wynia R, Thompson JA, Baker DW. 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]
6. 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]
7. Or CK, Karsh BT. A systematic review of patient acceptance of consumer health information technology. J Am Med Inform Assoc 2009 Aug;16(4):550-560 [FREE Full text] [doi: 10.1197/jamia.M2888] [Medline: 19390112]
8. Agarwal R, Anderson C, Zarate J, Ward C. If we offer it, will they accept? Factors affecting patient use intentions of personal health records and secure messaging. J Med Internet Res 2013;15(2):e43 [FREE Full text] [doi: 10.2196/jmir.2243] [Medline: 23470453]
9. Riippa I, Linna M, Rönkkö I. The Effect of a Patient Portal With Electronic Messaging on Patient Activation Among Chronically Ill Patients: Controlled Before-and-After Study. J Med Internet Res 2014;16(11):e257 [FREE Full text] [doi: 10.2196/jmir.3462] [Medline: 25413368]
10. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40(5):373-383. [Medline: 3558716]
11. 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]
12. McHorney CA, Ware JE, Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care 1993 Mar;31(3):247-263. [Medline: 8450681]
13. McHorney CA, Ware JE, Lu JF, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. Med Care 1994 Jan;32(1):40-66. [Medline: 8277801]
14. Hibbard JH, Mahoney ER, Stockard J, Tusler M. Development and testing of a short form of the patient activation measure. Health Serv Res 2005 Dec;40(6 Pt 1):1918-1930 [FREE Full text] [doi: 10.1111/j.1475-6773.2005.00438.x] [Medline: 16336556]
15. Stagg V. CHARLSON Stata module to calculate Charlson index of comorbidity. Boston College; 2006. URL: http://ideas.repec.org/c/boc/bocode/s456719.html [accessed 2014-07-14] [WebCite Cache ID 6R3rPih9b]
16. Earnest MA, Ross SE, Wittevrongel L, Moore LA, Lin CT. Use of a patient-accessible electronic medical record in a practice for congestive heart failure: patient and physician experiences. J Am Med Inform Assoc 2004;11(5):410-417 [FREE Full text] [doi: 10.1197/jamia.M1479] [Medline: 15187074]
17. Lau AY, Proudfoot J, Andrews A, Liaw ST, Crimmins J, Arguel A, et al. Which bundles of features in a Web-based personally controlled health management system are associated with consumer help-seeking behaviors for physical and emotional well-being? J Med Internet Res 2013;15(5):e79 [FREE Full text] [doi: 10.2196/jmir.2414] [Medline: 23649790]

XSL•FO
RenderX

18. 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]
19. Weingart SN, Rind D, Tofias Z, Sands DZ. Who uses the patient internet portal? The PatientSite experience. J Am Med Inform Assoc 2006;13(1):91-95 [FREE Full text] [doi: 10.1197/jamia.M1833] [Medline: 16221943]
20. Yamin CK, Emani S, Williams DH, Lipsitz SR, Karson AS, Wald JS, et al. The digital divide in adoption and use of a personal health record. Arch Intern Med 2011 Mar 28;171(6):568-574. [doi: 10.1001/archinternmed.2011.34] [Medline: 21444847]
21. Varsi C, Gammon D, Wibe T, Ruland CM. Patients' reported reasons for non-use of an internet-based patient-provider communication service: qualitative interview study. J Med Internet Res 2013;15(11):e246 [FREE Full text] [doi: 10.2196/jmir.2683] [Medline: 24220233]
22. Tenforde M, Nowacki A, Jain A, Hickner J. The association between personal health record use and diabetes quality measures. J Gen Intern Med 2012 Apr;27(4):420-424 [FREE Full text] [doi: 10.1007/s11606-011-1889-0] [Medline: 22005937]
23. Hibbard JH, Greene J. Who are we reaching through the patient portal: engaging the already engaged? Int J Pers Centered Med 2011;1(4):788-793.
24. Roblin DW, Houston TK, Allison JJ, Joski PJ, Becker ER. Disparities in use of a personal health record in a managed care organization. J Am Med Inform Assoc 2009 Oct;16(5):683-689 [FREE Full text] [doi: 10.1197/jamia.M3169] [Medline: 19567790]
25. Weppner WG, Ralston JD, Koepsell TD, Grothaus LC, Reid RJ, Jordan L, et al. Use of a shared medical record with secure messaging by older patients with diabetes. Diabetes Care 2010 Nov;33(11):2314-2319 [FREE Full text] [doi: 10.2337/dc10-1124] [Medline: 20739686]
26. Lustria ML, Smith SA, Hinnant CC. Exploring digital divides: an examination of eHealth technology use in health information seeking, communication and personal health information management in the USA. Health Informatics J 2011 Sep;17(3):224-243. [doi: 10.1177/1460458211414843] [Medline: 21937464]
27. Official Statistics of Finland (OSF). Use of information and communications technology by individuals e-publication URL: http://tilastokeskus.fi/til/sutivi/2013/sutivi_2013_2013-11-07_tie_001_en.html [accessed 2014-07-14] [WebCite Cache ID 6R3sCBIPC]
28. Rosenstock IM. Why people use health services. Milbank Mem Fund Q 1966 Jul;44(3):Suppl:94-Suppl127. [Medline: 5967464]
29. Grossman M. On the concept of health capital and the demand for health. J Polit Econ 1972;80(2):223-255.
30. Fries JF, Koop CE, Sokolov J, Beadle CE, Wright D. Beyond health promotion: reducing need and demand for medical care. Health Aff (Millwood) 1998 Apr;17(2):70-84 [FREE Full text] [Medline: 9558786]

Abbreviations

BMI: body mass index
CCI: Charlson comorbidity index
EHR: electronic health record
HbA1c: glycated hemoglobin
LDL: low-density lipoprotein
PAM: patient activation measure

Edited by G Eysenbach; submitted 22.07.14; peer-reviewed by D Gammon, T Palen, V Wolcott; comments to author 07.08.14; revised version received 24.10.14; accepted 03.11.14; published 08.12.14.

Please cite as:
Riippa I, Linna M, Rönkkö I, Kröger V
Use of an Electronic Patient Portal Among the Chronically Ill: An Observational Study
J Med Internet Res 2014;16(12):e275
URL: http://www.jmir.org/2014/12/e275/
doi:10.2196/jmir.3722
PMID:25488754

©Iiris Riippa, Miika Linna, Ilona Rönkkö, Virpi Kröger. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 08.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), which permits unrestricted use, distribution, and reproduction in any

XSL•FO
RenderX

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

Personal Health Record Reach in the Veterans Health Administration: A Cross-Sectional Analysis

Stephanie Leah Shimada[1,2,3], PhD; Cynthia A Brandt[4,5,6], MD, MPH; Hua Feng[1,3], MS; D Keith McInnes[1,2], MS, ScD; Sowmya R Rao[1,7], PhD; James A Rothendler[1,2], MD; David A Haggstrom[8,9], MAS, MD; Erica A Abel[4,5], PhD; Lisa S Cioffari[1,3], BA; Thomas K Houston[1,3], MD, MPH

[1]Center for Healthcare Organization and Implementation Research / eHealth Quality Enhancement Research Initiative, Edith Nourse Rogers Memorial VA Medical Center, Bedford, MA, United States

[2]Department of Health Policy and Management, Boston University School of Public Health, Boston, MA, United States

[3]Division of Health Informatics and Implementation Science, Department of Quantitative Health Sciences, University of Massachusetts Medical School, Worcester, MA, United States

[4]Pain Research, Informatics, Multi-morbidities, and Education Center, VA Connecticut Healthcare System, West Haven, CT, United States

[5]Yale Center for Medical Informatics, Yale School of Medicine, New Haven, CT, United States

[6]Department of Emergency Medicine, Yale School of Medicine, New Haven, CT, United States

[7]Division of Biostatistics and Health Services Research, Department of Quantitative Health Sciences, University of Massachusetts Medical School, Worcester, MA, United States

[8]Center for Health Information and Communication, Richard L. Roudebush VA Medical Center, Indianapolis, IN, United States

[9]Division of General Internal Medicine and Geriatrics, Department of Medicine, Indiana University School of Medicine, Indianapolis, IN, United States

Corresponding Author:
Stephanie Leah Shimada, PhD
Center for Healthcare Organization and Implementation Research / eHealth Quality Enhancement Research Initiative
Edith Nourse Rogers Memorial VA Medical Center
200 Springs Road (152)
Bedford, MA, 01730
United States
Phone: 1 781 687 2208
Fax: 1 781 687 3463
Email: stephanie.shimada@va.gov

Abstract

Background: My Health*e*Vet (MHV) is the personal health record and patient portal developed by the United States Veterans Health Administration (VA). While millions of American veterans have registered for MHV, little is known about how a patient's health status may affect adoption and use of the personal health record.

Objective: Our aim was to characterize the reach of the VA personal health record by clinical condition.

Methods: This was a cross-sectional analysis of all veterans nationwide with at least one inpatient admission or two outpatient visits between April 2010 and March 2012. We compared adoption (registration, authentication, opt-in to use secure messaging) and use (prescription refill and secure messaging) of MHV in April 2012 across 18 specific clinical conditions prevalent in and of high priority to the VA. We calculated predicted probabilities of adoption by condition using multivariable logistic regression models adjusting for sociodemographics, comorbidities, and clustering of patients within facilities.

Results: Among 6,012,875 veterans, 6.20% were women, 61.45% were Caucasian, and 26.31% resided in rural areas. The mean age was 63.3 years. Nationwide, 18.64% had registered for MHV, 11.06% refilled prescriptions via MHV, and 1.91% used secure messaging with their clinical providers. Results from the multivariable regression suggest that patients with HIV, hyperlipidemia, and spinal cord injury had the highest predicted probabilities of adoption, whereas those with schizophrenia/schizoaffective disorder, alcohol or drug abuse, and stroke had the lowest. Variation was observed across diagnoses in actual (unadjusted) adoption and use, with registration rates ranging from 29.19% of patients with traumatic brain injury to 14.18% of those with schizophrenia/schizoaffective disorder. Some of the variation in actual reach can be explained by facility-level differences in MHV adoption and by differences in patients' sociodemographic characteristics (eg, age, race, income) by diagnosis.

XSL•FO
RenderX

Conclusions: In this phase of early adoption, opportunities are being missed for those with specific medical conditions that require intensive treatment and self-management, which could be greatly supported by functions of a tethered personal health record.

(J Med Internet Res 2014;16(12):e272) doi:10.2196/jmir.3751

KEYWORDS

personal health records; patient characteristics; medical conditions; veterans

Introduction

The National Committee on Vital and Health Statistics recently recommended adopting the term personal health record (PHR) to refer to "the collection of information about an individual's health and health care, stored in electronic format" [1]. Although empiric evidence for the effectiveness of PHRs is limited, these systems have the potential to support the transformation of care from episodic visit-based care to continuous, coordinated care [2,3]. PHRs vary significantly in their functionality and content [1,4-6]. My Health*e*Vet (MHV), the Veterans Health Administration's PHR and patient portal, has multiple levels of access with increasing levels of functionality [7]. Anyone can access health education materials through the website. Those who register for an account can make use of personal health journals, track personal health care information and health measurements (ie, self-enter and track diet, activity, and vital signs), and set personal health care goals [8]. As the PHR is tethered to the US Department of Veterans Affairs (VA) electronic health record (EHR), registrants who are VA patients can also refill prescriptions online. Further, "authenticated" registrants who have also gone through an identity-verification process currently have access to additional EHR-extracted information including their problem lists, medication histories, laboratory results, appointment information, inpatient discharge summaries, and clinical notes. Authenticated patients can also send and receive secure messages with their primary care and specialist teams through MHV. Figure 1 is a screenshot of the MHV home page, showing the options available to users after they log in. New MHV features and functions are constantly being developed and are released periodically with software updates. The above describes the functionality of MHV at the end of 2012. As the PHR of the United States' largest integrated health care system, MHV has been adopted by over 2.3 million registrants and represents an ideal opportunity to study PHR systems [9].

Figure 1. View of MHV home page after login.



XSL•FO
RenderX

The potential of PHRs is dependent on the reach of these systems to patients with specific health problems who would most benefit from access to these functions. Reach—the proportion and characteristics of the target population that can access and is willing to use an innovation [10]—is an important concept for evaluation of new technologies [7]. As implementation of PHRs continues to increase, variations in reach of PHRs have been associated with patient characteristics such as age, income, and race [6,11]. Among veterans responding to the 2008 VA Survey of Healthcare Experiences of Patients (SHEP), use of MHV was associated with lower health status [12]. However, the 2010 National Survey of Veterans found that general health status did not significantly predict use of the PHR [13]. Outside the VA, there have also been contradictory reports regarding the association of health status and PHR use, with some studies suggesting that healthier patients with fewer prescriptions were more likely to use PHRs [14,15], while others indicated that sicker patients were more likely to use PHRs [11,16,17]. These contradictory findings suggest a complex relationship between health status and PHR adoption and use that may be dependent on patient characteristics, including demographics (eg, age, gender, race, income, urban vs rural location) and specific medical conditions. Previous articles have proposed that information seeking and engagement in patient technologies may be driven by uncertainty in diagnosis or treatment associated with the patients' condition, and perhaps condition-specific stigma [18]. Research on PHR adoption and use by patients with varied conditions may help clarify whether patient needs that are specific to each condition (eg, the need for frequent contact with providers, the need for self-management, or the desire to avoid stigma) may drive adoption and use. While there have been a number of studies examining barriers and facilitators of PHR use [19], and the effects of PHR use among groups of patients with specific clinical diagnoses such as diabetes [20] or mental health [21], little research has been conducted to understand differences in PHR reach by patients' specific clinical diagnoses.

We evaluated the reach of the VA PHR among 6 million US veterans nationwide actively receiving health care from the Veterans Health Administration. The primary goal was to conduct an exploratory analysis to improve our understanding of differences in PHR reach by examining adoption and use of MHV among patients with specific chronic trauma-related, mental, and medical health conditions. We compared the relative reach of MHV using a set of 18 health conditions of high prevalence and/or importance among veterans. We then further explored how patient demographics, comorbidities, and facility characteristics may be affecting MHV reach.

Methods

Study Design

We conducted a cross-sectional assessment of PHR adoption and use across clinical conditions. Use was assessed as of April 1, 2012. This study was approved by the Human Research Protection Committees at the VA Connecticut Healthcare System and the Yale School of Medicine, and the Institutional Review Board at the Edith Nourse Rogers VA Hospital in Bedford, Massachusetts.

Study Population

The study population included all American veterans age 18-100 years who had obtained care from the Veterans Health Administration between April 1, 2010, and March 31, 2012 (N=6,012,875). Cohort inclusion criteria included at least two outpatient visits or one inpatient hospitalization for any cause during this period.

Data

Overview

We used data from the VA system of records available through the VA Corporate Data Warehouse. Variables from the Electronic Health Record included patient demographics and International Classification of Disease (ICD-9-CM) diagnosis codes associated with all VA inpatient and outpatient encounters from October 1, 2007, to March 31, 2012. These data were linked at the patient level with MHV registration, authentication, secure messaging (SM) opt-in, and use log data from April 2012.

Definition of Health Conditions (Independent Variables)

We defined patients as having medical and mental health conditions that were coded with ICD-9-CM codes at least once for an inpatient stay or at least twice for outpatient visits. We used previously validated diagnostic code groupings [22] to identify patients with one or more high prevalence or high priority clinical conditions. We selected 18 conditions based on their being high prevalence within the VA population and/or because they were one of the conditions that are high priority for VA clinicians and researchers as evidenced by their inclusion in the VA Quality Enhancement Research Initiative (QUERI) Program, for example, traumatic brain injury (TBI), spinal cord injury (SCI), and human immunodeficiency virus (HIV). Because outpatient codes are commonly assigned by health care providers in the VA and tend to be less accurate than inpatient codes assigned by professional coders, we combined inpatient and outpatient data sources. This has been shown to improve the accuracy of identifying psychiatric disorders [23] and HIV in administrative data [24]. The 18 conditions represented trauma-related (SCI, TBI), mental health (common mood disorders, eg, mild and major depression, anxiety, and posttraumatic stress disorder), substance use disorders (alcohol abuse, drug abuse), other serious mental illnesses (schizophrenia, psychoses), and medical (chronic heart failure, diabetes, hyperlipidemia, hypertension, HIV infection, hepatitis, ischemic heart disease, stroke) clinical conditions.

Measuring My HealtheVet Reach (Dependent Variables)

Our measures of MHV reach included indicators for adoption of the PHR as of April 2012: (1) registered MHV user ("registered"), (2) registered and in-person authenticated ("authenticated"), and (3) authenticated and opted-in to use secure messaging ("SM opt-in"). Further, we calculated indicators for use of two core PHR functions: (1) use of the online prescription refill feature ("refills") in MHV, and (2) sending any secure messages to providers ("messages sent").

XSL•FO
RenderX

Sociodemographic Characteristics (Covariates)

Variables were available for the following sociodemographic characteristics: age, gender, race/ethnicity (white, African-American, Latino, other, and unknown), urban/rural residence based on home postal code, and economic need defined as eligibility for free care based on an annual VA financial assessment.

Statistical Analyses

Sociodemographic Characteristics

We first compared the sociodemographic characteristics of veterans across the 18 specific clinical conditions. For sociodemographic characteristics, rates of missing were 0.2% for age, 0.02% for gender, 2.2% for means test, and 4.1% for urban/rural. Only race/ethnicity had a missing data rate of larger than 5%. We categorized the 18.38% of patients with unknown (to patient) or missing race into a group (unknown/missing) and included them in our analysis but excluded patients with missing values on other variables.

Main Analysis: Variation of Personal Health Record Use by Clinical Condition

For our main analysis, we assessed reach (adoption and use) of the PHR across the 18 specific clinical conditions. As noted, our main dependent variables were PHR adoption and use. Our main independent variables were the 18 specific clinical conditions. We calculated means, standard deviations, and distributions of veterans' demographic characteristics, PHR adoption tier (registered, authenticated, opted-in to SM), and indicators of use (messages sent, prescription refills) in all subjects and by clinical condition. To understand the current relative use of the PHR, we then conducted bivariate, unadjusted analyses of the dichotomous MHV adoption and use variables by each condition.

To further understand how complex variations in patient characteristics and facility-level differences in MHV adoption might bias the primary association of specific clinical conditions and MHV adoption, we then obtained adjusted predicted probabilities (adjusted for age, gender, race/ethnicity, economic need via means test, and urban/rural status) and 95% confidence intervals from multivariable models, accounting for clustering of veterans within facilities and including facility as a random effect to adjust for facility-level differences. We ran our multivariable logistic regression models using generalized linear mixed models with a binomial distribution and logit link. We converted the least squares means obtained for each condition from these models into the predicted probabilities (shown as percentages in the tables) using the ILINK option. We also calculated the intraclass correlation coefficient in order to evaluate the potential impact of facility-level variation on the patient-level associations with PHR adoption. A two-sided $P<.05$ was considered to be significant. All analyses were conducted using SAS and SAS Grid 9.2.

To better visualize the result of adjustment for sociodemographic characteristics on the primary association of authentication and clinical condition, we created a multi-attribute plot. This plot shows the association between patient age and economic need, and pre- to post-adjustment change in the relative rank of conditions based on predicted percentages of patients with each condition adopting the PHR. Increasing age and higher economic need were selected for visualization as they are known from prior studies to be aspects of the "digital divide" associated with lower PHR use [11,25].

Results

Clinical and Sociodemographic Characteristics

Out of over 6 million (6,012,875) VA patients nationwide, 4,893,286 (81.38%) had one or more of our 18 target conditions, and 1,119,589 (18.62%) had none of the 18 conditions. The most prevalent specific clinical conditions were hypertension (56.63% of patients), hyperlipidemia (55.69%), diabetes (24.71%), and depression (24.68%). The least prevalent conditions were spinal cord injury (0.43%), HIV (0.45%), and TBI (1.40%). Overall, the population had a mean age of 63.3 years and was 6.20% female, 61.45% white, 73.69% urban, and 26.76% were eligible for free care based on a VA financial assessment (see Table 1). Patients without any of the conditions were younger, more likely to be female, urban residents, of unknown race, and less likely to be eligible for free care than patients with at least one condition. See Multimedia Appendix 1 for a table showing demographics by clinical condition.

XSL•FO
RenderX

Table 1. Demographic characteristics of patients, overall and by presence of conditions.

Characteristics	Overall (N=6,012,875)	Patients with at least one of 18 conditions[a] (N=4,893,286)	Patients without any of 18 conditions (N=1,119,589)
Age, mean (SD)	63.32 (16.46)	65.20 (15.02)	55.12 (19.66)
Gender, %			
Female	6.20	5.33	9.98
Male	93.80	94.67	90.02
Race, %			
White	61.45	62.55	56.66
African-American	13.09	13.27	12.28
Latino	5.46	5.39	5.77
Other	1.62	1.49	2.20
Unknown/Missing	18.38	17.30	23.09
High economic need[b], %	26.76	27.30	24.34
Urban/Rural, %			
Urban	73.69	72.73	77.95
Rural	26.31	27.27	22.05

[a]18 conditions: Medical—hypertension, hyperlipidemia, diabetes, coronary artery disease, chronic obstructive pulmonary disease, congestive heart failure, stroke, hepatitis, HIV; Mental health—anxiety, depression, post-traumatic stress disorder, psychosis, schizophrenia/schizoaffective disorder, alcohol abuse, drug abuse; Trauma-related—traumatic brain injury, spinal cord injury.

[b]Eligible for free care based on a VA financial assessment.

Variation of Reach (Personal Health Record Adoption and Use) by Clinical Condition

As of April 2012, reach of the PHR remained relatively low. Among patients who had been seen in VA between April 2010 and March 2012, 1.12 million (18.64%) were registered, 0.6 million (10.03%) of these were authenticated, 0.24 million (4.05%) had opted in for secure messaging, and 0.67 million (11.06%) of registered veterans had used the PHR for a prescription refill. Of those opted-in to secure messaging, 47.19% had sent at least one message to their clinical team since opting in. There was significant variation across facilities in PHR adoption. Registration rates varied from 9.30% to 34.91% across facilities (mean 18.82%, SD 5.02), and authentication rates varied from 3.44% to 30.39% (mean 10.76%, SD 4.84).

Reach varied significantly by condition. Table 2 shows the unadjusted breakdown of PHR adoption and use by specific clinical conditions, reflecting actual adoption and use across the VA. Unadjusted adoption was generally higher among patients with trauma-related diagnoses, mood disorders, and posttraumatic stress disorder (PTSD). Veterans with PTSD, TBI, spinal cord injury, depression, anxiety, and HIV more frequently adopted the PHR, compared to those with other conditions. Patients with complex, chronic medical conditions such as hepatitis, coronary artery disease, congestive heart failure, or schizophrenia were less likely to have adopted the PHR.

Higher rates of adoption were also associated with higher levels of use. Among patients with PTSD, 17.01% were authenticated, 18.04% had refilled a medication through the PHR, and 7.00% had opted in to secure messaging, with 3.41% having actually sent a secure message by April 2012. In contrast, 8.05% of patients with schizophrenia/schizoaffective disorder were authenticated, only 8.09% had refilled a prescription through the PHR, and only 1.11% had sent a message. Generally, the reach of the PHR was lower across chronic, complex medical conditions relative to those with trauma-related or mental health diagnoses. Among veterans with coronary artery disease, 11.11% were refilling medications through the PHR, and 1.85% secure messaging, and use was even lower among those with congestive heart failure (10.38% and 1.69%, refills and secure messaging respectively). Those without any of the 18 conditions were considerably less likely than those with one or more diagnoses to have adopted MHV or to have used either of its key features.

XSL•FO
RenderX

Table 2. My Health*e*Vet reach (adoption and use) by specific clinical condition[a].

Condition group (N=6,012,875)	Registered, % (N=1,120,667)	Authenticated, % (N=603,054)	Opted into secure messaging, % (N=243,456)	Sent at least one secure message, % (N=114,884)	Used medication refill, % (N=665,291)
Overall	18.64	10.03	4.05	1.91	11.06
Traumatic brain injury	29.19	16.53	6.39	3.24	17.69
PTSD	28.42	17.01	7.00	3.41	18.04
Spinal cord injury	27.37	16.29	6.65	3.05	18.84
HIV	26.48	15.31	6.30	2.41	17.76
Anxiety	26.39	15.37	6.38	3.16	16.87
Depression	26.17	15.33	6.35	3.15	16.80
Diabetes	21.17	12.52	5.24	2.48	13.60
Hyperlipidemia	20.91	11.76	4.92	2.31	13.20
Psychosis	20.29	11.80	4.38	2.16	12.40
Hypertension	19.70	11.10	4.62	2.17	12.36
Alcohol abuse	19.38	11.18	4.08	1.88	11.38
Drug abuse	18.90	11.21	3.72	1.73	10.32
Chronic obstructive pulmonary disease	18.86	11.04	4.43	2.07	11.86
Hepatitis	18.04	10.89	4.10	1.89	10.16
Coronary artery disease	17.95	10.17	4.11	1.85	11.11
Congestive heart failure	17.22	9.95	3.77	1.69	10.38
Stroke	17.15	9.73	3.68	1.70	10.35
Schizophrenia or schizoaffective disorder	14.18	8.05	2.61	1.11	8.09
None of the diagnoses above	10.81	4.81	1.62	0.73	4.33

[a]These are actual percentages of patients in each condition group, unadjusted for clustering.

Multivariable Adjustment

Our analyses to explore whether the patient-level findings were influenced by the facility-level adoption of MHV showed significant variation at the facility level (P<.001), but this variation did not alter the associations between diagnoses and MHV adoption found at the patient level. Rank ordering of the conditions based on adoption remained largely unchanged after inclusion of facility as a random effect. We found that only 13.7% of the measured variance was at the facility-level, suggesting that the majority of the differences in adoption are predicted by patient characteristics. After inclusion of sociodemographic characteristics and comorbidities, the rank ordering of conditions and the predicted percentage of patients authenticated for MHV use did change (Table 3). HIV, hyperlipidemia, and SCI are the three conditions with the highest predicted percentages of authentication, and both hyperlipidemia and hypertension move up in rank to be among the conditions with the highest predicted percentage of authentication. Younger patients, white patients, women, and patients who did not qualify for free care based on financial need were more likely to be authenticated (results not shown). All sociodemographic variables were significant predictors of authentication; however, age and financial need had the greatest impact on adjusted rates. Authentication rates were adjusted downwards for conditions where patients were younger than average (eg, patients with TBI, PTSD), and upward for conditions where patients tend to be older than average (eg, stroke, coronary artery disease). These effects are mitigated or reversed where a larger-than-average percentage of the population is higher in financial need.

XSL•FO
RenderX

Table 3. Unadjusted and adjusted predicted percentages of authentication by specific clinical condition.

Condition group (N=6,012,875)[a]	Unadjusted predicted percentages of authentication (CI)	Adjusted[b] predicted percentages of authentication (CI)	Difference (adjusted–unadjusted)
HIV	15.33 (14.26-16.46)	16.20 (15.01-17.46)	0.87
Hyperlipidemia	11.78 (10.98-12.64)	16.11 (14.99-17.30)	4.33
Spinal cord injury	16.36 (15.23-17.55)	15.74 (14.58-16.98)	-0.62
Depression	15.21 (14.23-16.26)	15.34 (14.26-16.47)	0.13
Hypertension	11.05 (10.30-11.84)	14.97 (13.92-16.09)	3.92
Post-traumatic stress disorder	16.60 (15.56-17.70)	14.66 (13.62-15.76)	-1.94
Diabetes	12.51 (11.68-13.39)	14.53 (13.50-15.62)	2.02
Anxiety	15.37 (14.37-16.41)	13.95 (12.96-15.01)	-1.42
Traumatic brain injury	16.62 (15.55-17.76)	13.77 (12.77-14.83)	-2.85
Chronic obstructive pulmonary disease	11.08 (10.34-11.88)	13.47 (12.51-14.50)	2.39
Hepatitis	10.64 (9.91-11.41)	13.34 (12.38-14.36)	2.70
Psychosis	11.63 (10.85-12.46)	13.02 (12.09-14.02)	1.39
Coronary artery disease	10.26 (9.56-11.00)	12.78 (11.86-13.76)	2.52
Congestive heart failure	9.89 (9.21-10.61)	12.64 (11.72-13.62)	2.75
Stroke	9.64 (8.97-10.35)	12.37 (11.46-13.33)	2.73
Drug abuse	11.04 (10.29-11.83)	12.29 (11.39-13.24)	1.25
Alcohol abuse	10.90 (10.16-11.68)	12.08 (11.20-13.02)	1.18
Schizophrenia / schizoaffective disorder	8.04 (7.47-8.65)	10.40 (9.62-11.25)	2.36

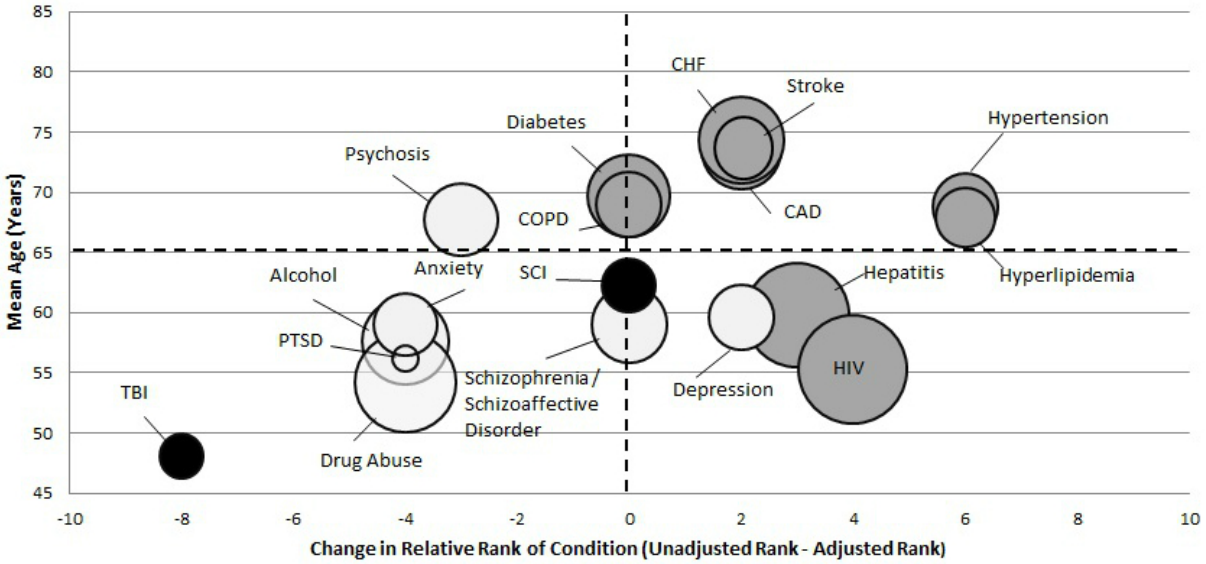
[a]After excluding patients with missing data for facility or one of the covariates, data for 5,988,043 observations were used for logistic regression analysis. Facility was included as a random effect in both unadjusted and adjusted logistic regression models to adjust for clustering and IPA variation between facilities. For all disease conditions, $P<.001$ in both unadjusted and adjusted logistic regression analyses, except congestive heart failure ($P=.74$) in the unadjusted model.

[b]Adjusted for presence of each other specific clinical condition and demographics including age (continuous), gender, race/ethnicity, economic need, and rural/urban residence. Race/ethnicity includes non-Hispanic white, non-Hispanic African-American, hispanic, other, and unknown/missing with white as reference group.

To better visualize effect modification of the association of authentication and clinical condition by age and economic need, we created a multi-attribute plot (see Figure 2). The horizontal axis shows the change in the relative rank after adjustment. The vertical axis is the mean age of veterans in the disease category. Further, the size of the bubbles is proportional to the percent of the population qualifying for free care based on income. The shading of the bubbles distinguishes trauma-related (black), mental health (white), and medical conditions (gray). Using dotted lines, we divided the figure into four quadrants divided by age (above vs below the mean age) and change in rank (above zero=increase in rank after adjustment). For example, we see in the upper right quadrant that all medical conditions affecting older than average patient populations (coronary artery disease, congestive heart failure, chronic obstructive pulmonary disease, stroke, diabetes, hypertension, and hyperlipidemia) have increased in rank post-adjustment.

XSL•FO
RenderX

Figure 2. Bubble plot of conditions by mean age (years), change in rank, and socioeconomic status. CAD: coronary heart disease; CHF: congestive heart failure; COPD: chronic obstructive pulmonary disease; HIV: human immunodeficiency virus; PTSD: post-traumatic stress disorder; SCI: spinal cord injury; TBI: traumatic brain injury.



Discussion

Principal Results

As of April 2012, the reach of My Health*e*Vet, a national, multifunction, tethered PHR was less than 20% of veterans actively receiving care in VA, and no clinical group had adoption over 30%. Reach had increased from 16.3% in July 2009, as reported by Nazi et al [7,26] and was slightly lower than the 21.51% estimated by Tsai and Rosenheck [13] based on a survey of 7215 veterans conducted in 2010. During this early phase of adoption, we found that patients with trauma-related conditions and common mental health conditions were among the highest adopters, and patients with certain chronic, complex medical conditions were lower adopters.

Adjusting for sociodemographic characteristics and comorbidities had a significant impact on the predicted probability of authentication. In April 2012, patients with PTSD, TBI, and SCI were the highest adopters of MHV. However, after adjustment, patients with HIV, hyperlipidemia, and SCI were predicted to be the most likely to authenticate. These changes illustrate that differences in actual adoption and use across conditions are at least partially driven by differences in the sociodemographic characteristics of patients by condition. In prior studies, age, ethnicity, and income have been identified as important patient characteristics that predict Internet use in general, and patient portal use more specifically [27]. There were also large facility-level differences in the level of MHV adoption achieved, suggesting that facilities vary meaningfully in the outreach and support they provide to patients regarding adoption and use of MHV. Adjusting for sociodemographic characteristics, comorbidities, and facility-level differences is therefore important when studying PHR reach.

We did find that differences in adoption and use of MHV by health condition remained after adjustment for sociodemographic factors, comorbidities, and facility level variation. All conditions remained significantly associated with authentication in our multivariate model ($P<.001$ for all), and there were statistically significant differences across conditions in the predicted percentages of patients authenticating to MHV. How meaningful those differences are will be determined by the availability of PHR-based tools and interventions to improve self-management and/or health outcomes for patients with that condition. At the same time, the feasibility of developing and implementing such tools and interventions will likely be driven by PHR reach to patients with that condition.

Veterans with trauma-related conditions (TBI and SCI) were frequent PHR users in April 2012. After adjustment, only patients with SCI continued to have a high predicted percentage of authentication. The increased adoption among patients with TBI was largely moderated by their younger age (see Figure 1). Diagnosis of TBI continues to increase [28,29] among young veterans of recent conflicts. The cohort of veterans with TBI is younger, and the relative higher prevalence of adoption, as compared with the other specific clinical conditions, is highly moderated by age. To maximize the effectiveness of the PHR for this group of patients, an adaptive PHR user interface that specifically addresses the cognitive issues in this population may be required. Previous articles have proposed that information seeking and engagement in patient technologies may be driven by uncertainty in diagnosis or treatment associated with the patients' condition [18]. Because TBI is a condition with varied presentation and uncertain treatment and outcomes, these proposed forces may be driving increased utilization.

Reach of the PHR among patients with depression, anxiety, and PTSD was also relatively high in April 2012. Although a previous study did not detect a significant overall difference in use between veterans receiving care in the VA with and without mental health diagnoses [13], our larger sample size has enabled us to detect differences in actual adoption and use. While these differences are attenuated after adjusting for sociodemographic

XSL•FO
RenderX

characteristics, comorbidities, and clustering of patients within facilities, having one of the above mental health diagnoses continues to be a significant predictor of authentication. Patients with depression, anxiety, and PTSD also have higher levels of use of PHRs. A relatively high percentage of these patients had used online prescription refills, and they were among the most likely to have used secure messaging. This was true despite the fact that secure messaging had not yet been implemented among mental health providers and was largely limited to primary care in April 2012. Depression, anxiety, and PTSD are prevalent in the general patient population as well as in VA [28,30-37]. Health care providers have expressed concern that mental health patients would be frequent users of the system, overloading the clinical team with questions and potentially using the system inappropriately. Our results suggest that there may be some reality underlying provider perceptions that patients with specific mental health conditions will frequently use the system. This may be due in part to the fact that patients with multiple medical conditions are more likely to also have mental health conditions [38], thus requiring more health care overall. Future research should address the effectiveness of mental health care via asynchronous secure messaging to assess whether PHRs and other patient-facing technologies improve appropriate access to care or increase inappropriate use.

Among medical conditions, patients with HIV had the highest levels of adoption of the PHR in April 2012, followed by those with diabetes. After adjustment for demographic characteristics, comorbid conditions, and for clustering of patients within facilities, patients with HIV, hyperlipidemia, hypertension, and diabetes had high probabilities of authentication. These are complex conditions, requiring multiple medications and frequent medical follow-up visits. We speculate that higher service utilization and increased need for self-management in these conditions are likely driving PHR adoption. The influence between patient technologies and service utilization may be bi-directional. Patients with more frequent visits to the VA may be more likely to be exposed to information on MHV and encouraged to register and authenticate by their providers and clinic staff. Introduction to PHR functions such as secure messaging can also result in increased service utilization, perhaps by making it easier for patients to raise concerns via online communication with their providers [17]. These conditions also benefit from better nutrition and medication management, as well as improved tracking and monitoring of vitals and readings (eg, blood pressure, blood sugar, weight) and viewing of chemistry/hematology laboratory results (eg, hemoglobin A1c, lipids, CD4+ cell counts, viral load), all of which can be accomplished with assistance from the PHR. In addition, the more medications a patient has, the more potential benefit they may derive from access to the medication refill function of the PHR.

Limitations

Because the veteran population differs from the general US population in many ways, including higher economic need, higher burden of substance abuse and mental illness, and higher proportion of male patients [29,32,39], these results are not strictly generalizable to non-VA populations. However, it seems likely that adoption and use of PHRs is also driven by clinical need and moderated by patient sociodemographics in non-VA populations [16].

This analysis is limited in the types of PHR use we were able to measure. Understanding the extent of use of other MHV features currently available to veterans, such as medication management tools and wellness reminders, would be of future interest since these features have the potential to improve both medication adherence and evidence-based care [40-42]. In addition, knowledge regarding use of the chronic disease self-management component of MHV is of interest since such features have the potential to significantly improve health outcomes [43-45]. However, at the time of this analysis, our only available measures of "use" were the use of secure messaging or prescription refills. More detailed measures of use are needed to assess the impact of the PHR implementation on outcomes relevant to specific clinical conditions.

After adjusting for patients' sociodemographic characteristics, their other comorbidities, and the facility at which they received their care, we continued to observe differences in adoption and use of the PHR by diagnosis. However, these analyses were not able to uncover what drives these differences. Also, our analyses also did not focus on patterns of adoption and use among patients with common and/or costly combinations of chronic conditions (eg, diabetes, hyperlipidemia, and hypertension) [46]. Future research should attempt to understand drivers of and barriers to PHR adoption and use.

Implications for Future Research and PHR Interventions

Variations in adoption and use by diagnosis have implications for delivery of interventions through a PHR or patient portal. Although adoption remained low (less than 30% registered and 17% authenticated in April 2012) for each specific clinical condition, certain groups, such as younger patients with mental illness, may be more ready for and receptive to targeted interventions delivered through a PHR. Understanding the level of adoption and the types of use among patients with the most prevalent clinical conditions can help with prioritizing the development of eHealth tools with the potential to improve self-management and further engage a given patient subgroup.

Our results identified significant gaps in adoption and use. Specifically, reach among patients with certain complex, chronic medical conditions was lower than for those in the high-adoption conditions. Thus, in this phase of early adoption, opportunities are being missed for supporting those with medical conditions that require intensive treatment and self-management via functions of a tethered PHR. These groups should be provided with outreach and supported with interventions including training or educational materials and proactive "help desk" support.

Conclusions

To our knowledge, this is the first report using national EHR data to associate PHR reach with patient diagnoses. We reported unadjusted reach and then calculated adjusted predicted percentages. Both the unadjusted and adjusted measures have important implications. The unadjusted relative reach reflects the current reality within the Veterans Health Administration

XSL•FO
RenderX

and suggests that patients with specific clinical conditions may require additional interventions to support adoption and use. The adjusted estimates provide insights as to the effect of age and economic status on PHR adoption. Although the digital divide of technology access frequently cited in the literature [47] has narrowed, it still exists for older veterans [25] and those of higher economic need [48].

Although an earlier pilot version of My Health*e*Vet had high satisfaction [26] and appeared to improve patient-provider communication and patient engagement in care [49], considerable research needs to be done on the ability of PHRs to support continuous, coordinated, patient-centered, efficient care that is high quality and safe [7]. By first developing interventions for patient populations most ready to adopt them, while providing training and outreach to those groups lagging in adoption, we can move this research agenda forward more rapidly and effectively. Because many patients, both within VA and outside, are multimorbid with mental and medical conditions, future research should also examine how the need to manage multiple comorbid medical and mental health conditions impacts PHR adoption and use.

Acknowledgments

The research reported here was supported by the Department of Veterans Affairs (VA) Health Services Research and Development Service, Quality Enhancement Research Initiative (RRP 11-404). In addition, Dr Shimada was supported by a VA Health Services Research and Development Career Development Award (CDA# 10-210). The views expressed in this paper are those of the authors and do not necessarily represent the views of the Department of Veterans Affairs.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Demographic characteristics by condition, ranked by frequency.

[PDF File (Adobe PDF File), 13KB - jmir_v16i12e272_app1.pdf]

References

1. US Department of Health and Human Services. The National Infrastructure Workgroup of the National Committee on Vital Statistics (NICVHS). 2006. Personal Health Records and Personal Health Record Systems URL: http://www.ncvhs.hhs.gov/0602nhiirpt.pdf [accessed 2014-11-26] [WebCite Cache ID 6UNKGcoZW]
2. Committee on Quality of Health Care in America. Crossing the Quality Chasm. Washington DC: National Academy Press; 2001:164-206.
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. Archer N, Fevrier-Thomas U, Lokker C, McKibbon KA, Straus SE. Personal health records: a scoping review. J Am Med Inform Assoc 2011;18(4):515-522 [FREE Full text] [doi: 10.1136/amiajnl-2011-000105] [Medline: 21672914]
5. Detmer D, Bloomrosen M, Raymond B, Tang P. Integrated personal health records: transformative tools for consumer-centric care. BMC Med Inform Decis Mak 2008;8:45 [FREE Full text] [doi: 10.1186/1472-6947-8-45] [Medline: 18837999]
6. 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]
7. Nazi KM, Hogan TP, Wagner TH, McInnes DK, Smith BM, Haggstrom D, et al. Embracing a health services research perspective on personal health records: lessons learned from the VA My HealtheVet system. J Gen Intern Med 2010 Jan;25 Suppl 1:62-67 [FREE Full text] [doi: 10.1007/s11606-009-1114-6] [Medline: 20077154]
8. Department of Veterans Affairs. My HealtheVet Account Types. 2013 Mar 20. URL: https://www.myhealth.va.gov/mhv-portal-web/anonymous.portal?_nfpb=true&_nfto=false&_pageLabel=spotlightArchive&contentPage=ipa/mhv_account_status-definitions.html&WT.ac=splash_account_types
9. Shimada SL, Hogan TP, Rao SR, Allison JJ, Quill AL, Feng H, et al. Patient-provider secure messaging in VA: variations in adoption and association with urgent care utilization. Med Care 2013 Mar;51(3 Suppl 1):S21-S28. [doi: 10.1097/MLR.0b013e3182780917] [Medline: 23407007]
10. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am J Public Health 1999 Sep;89(9):1322-1327. [Medline: 10474547]
11. Yamin CK, Emani S, Williams DH, Lipsitz SR, Karson AS, Wald JS, et al. The digital divide in adoption and use of a personal health record. Arch Intern Med 2011 Mar 28;171(6):568-574. [doi: 10.1001/archinternmed.2011.34] [Medline: 21444847]

XSL•FO
RenderX

12. Veterans Health Administration Office of Health Information Product Effectiveness. My HealtheVet SHEP Survey Analysis Report. 2009. Evaluation of Quality Management in Veterans Health Administration Facilities URL: http://www.va.gov/oig/pubs/VAOIG-12-01480-183.pdf [accessed 2014-11-26] [WebCite Cache ID 6UNKiLjL1]
13. Tsai J, Rosenheck RA. Use of the internet and an online personal health record system by US veterans: comparison of Veterans Affairs mental health service users and other veterans nationally. J Am Med Inform Assoc 2012;19(6):1089-1094 [FREE Full text] [doi: 10.1136/amiajnl-2012-000971] [Medline: 22847305]
14. Weingart SN, Hamrick HE, Tutkus S, Carbo A, Sands DZ, Tess A, et al. Medication safety messages for patients via the web portal: the MedCheck intervention. Int J Med Inform 2008 Mar;77(3):161-168. [doi: 10.1016/j.ijmedinf.2007.04.007] [Medline: 17581772]
15. Weingart SN, Rind D, Tofias Z, Sands DZ. Who uses the patient internet portal? The PatientSite experience. J Am Med Inform Assoc 2006;13(1):91-95 [FREE Full text] [doi: 10.1197/jamia.M1833] [Medline: 16221943]
16. 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]
17. Palen TE, Ross C, Powers JD, Xu S. Association of online patient access to clinicians and medical records with use of clinical services. JAMA 2012 Nov 21;308(19):2012-2019. [doi: 10.1001/jama.2012.14126] [Medline: 23168824]
18. Davison KP, Pennebaker JW, Dickerson SS. Who talks? The social psychology of illness support groups. Am Psychol 2000 Feb;55(2):205-217. [Medline: 10717968]
19. Chumbler NR, Haggstrom D, Saleem JJ. Implementation of health information technology in Veterans Health Administration to support transformational change: telehealth and personal health records. Med Care 2011 Dec;49 Suppl:S36-S42. [doi: 10.1097/MLR.0b013e3181d558f9] [Medline: 20421829]
20. 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]
21. Druss BG, Ji X, Glick G, von Esenwein SA. Randomized trial of an electronic personal health record for patients with serious mental illnesses. Am J Psychiatry 2014 Mar;171(3):360-368. [doi: 10.1176/appi.ajp.2013.13070913] [Medline: 24435025]
22. Goulet JL, Fultz SL, Rimland D, Butt A, Gibert C, Rodriguez-Barradas M, et al. Aging and infectious diseases: do patterns of comorbidity vary by HIV status, age, and HIV severity? Clin Infect Dis 2007 Dec 15;45(12):1593-1601 [FREE Full text] [doi: 10.1086/523577] [Medline: 18190322]
23. Lurie N, Popkin M, Dysken M, Moscovice I, Finch M. Accuracy of diagnoses of schizophrenia in Medicaid claims. Hosp Community Psychiatry 1992 Jan;43(1):69-71. [Medline: 1544654]
24. Walkup JT, Wei W, Sambamoorthi U, Crystal S. Sensitivity of an AIDS case-finding algorithm: who are we missing? Med Care 2004 Aug;42(8):756-763. [Medline: 15258477]
25. Hall AK, Bernhardt JM, Dodd V, Vollrath MW. The Digital Health Divide: Evaluating Online Health Information Access and Use Among Older Adults. Health Educ Behav 2014 Aug 25:-. [doi: 10.1177/1090198114547815] [Medline: 25156311]
26. 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]
27. Bali RK, Gibbons MC. Exploring the link between populomics and knowledge management: patient-oriented case exemplars. Stud Health Technol Inform 2010;156:138-149. [Medline: 20543349]
28. Golub A, Vazan P, Bennett AS, Liberty HJ. Unmet need for treatment of substance use disorders and serious psychological distress among veterans: a nationwide analysis using the NSDUH. Mil Med 2013 Jan;178(1):107-114 [FREE Full text] [Medline: 23356128]
29. Lew HL, Otis JD, Tun C, Kerns RD, Clark ME, Cifu DX. Prevalence of chronic pain, posttraumatic stress disorder, and persistent postconcussive symptoms in OIF/OEF veterans: polytrauma clinical triad. J Rehabil Res Dev 2009;46(6):697-702 [FREE Full text] [Medline: 20104399]
30. Jonas BS, Brody D, Roper M, Narrow WE. Prevalence of mood disorders in a national sample of young American adults. Soc Psychiatry Psychiatr Epidemiol 2003 Nov;38(11):618-624. [doi: 10.1007/s00127-003-0682-8] [Medline: 14614549]
31. Ansseau M, Dierick M, Buntinkx F, Cnockaert P, De Smedt J, Van Den Haute M, et al. High prevalence of mental disorders in primary care. J Affect Disord 2004 Jan;78(1):49-55. [Medline: 14672796]
32. Magruder KM, Frueh BC, Knapp RG, Davis L, Hamner MB, Martin RH, et al. Prevalence of posttraumatic stress disorder in Veterans Affairs primary care clinics. Gen Hosp Psychiatry 2005;27(3):169-179. [doi: 10.1016/j.genhosppsych.2004.11.001] [Medline: 15882763]
33. Frueh BC, Grubaugh AL, Acierno R, Elhai JD, Cain G, Magruder KM. Age differences in posttraumatic stress disorder, psychiatric disorders, and healthcare service use among veterans in Veterans Affairs primary care clinics. Am J Geriatr Psychiatry 2007 Aug;15(8):660-672. [doi: 10.1097/JGP.0b013e3180487cc2] [Medline: 17670996]
34. Riolo SA, Nguyen TA, Greden JF, King CA. Prevalence of depression by race/ethnicity: findings from the National Health and Nutrition Examination Survey III. Am J Public Health 2005 Jun;95(6):998-1000. [doi: 10.2105/AJPH.2004.047225] [Medline: 15914823]

XSL•FO
RenderX

35. Grubaugh AL, Monnier J, Magruder KM, Knapp RG, Frueh BC. Female veterans seeking medical care at Veterans Affairs primary care clinics: psychiatric and medical illness burden and service use. Women Health 2006;43(3):41-62. [doi: 10.1300/J013v43n03_03] [Medline: 17194677]
36. Cohen BE, Gima K, Bertenthal D, Kim S, Marmar CR, Seal KH. Mental health diagnoses and utilization of VA non-mental health medical services among returning Iraq and Afghanistan veterans. J Gen Intern Med 2010 Jan;25(1):18-24 [FREE Full text] [doi: 10.1007/s11606-009-1117-3] [Medline: 19787409]
37. Seal KH, Maguen S, Cohen B, Gima KS, Metzler TJ, Ren L, et al. VA mental health services utilization in Iraq and Afghanistan veterans in the first year of receiving new mental health diagnoses. J Trauma Stress 2010 Feb;23(1):5-16. [doi: 10.1002/jts.20493] [Medline: 20146392]
38. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. Lancet 2012 Jul 7;380(9836):37-43 [FREE Full text] [doi: 10.1016/S0140-6736(12)60240-2] [Medline: 22579043]
39. Mattocks KM, Skanderson M, Goulet JL, Brandt C, Womack J, Krebs E, et al. Pregnancy and mental health among women veterans returning from Iraq and Afghanistan. J Womens Health (Larchmt) 2010 Dec;19(12):2159-2166 [FREE Full text] [doi: 10.1089/jwh.2009.1892] [Medline: 21039234]
40. Stone EG, Morton SC, Hulscher ME, Maglione MA, Roth EA, Grimshaw JM, et al. Interventions that increase use of adult immunization and cancer screening services: a meta-analysis. Ann Intern Med 2002 May 7;136(9):641-651. [Medline: 11992299]
41. Ayanian JZ, Sequist TD, Zaslavsky AM, Johannes RS. Physician reminders to promote surveillance colonoscopy for colorectal adenomas: a randomized controlled trial. J Gen Intern Med 2008 Jun;23(6):762-767 [FREE Full text] [doi: 10.1007/s11606-008-0576-2] [Medline: 18386103]
42. Sequist TD, Zaslavsky AM, Marshall R, Fletcher RH, Ayanian JZ. Patient and physician reminders to promote colorectal cancer screening: a randomized controlled trial. Arch Intern Med 2009 Feb 23;169(4):364-371 [FREE Full text] [doi: 10.1001/archinternmed.2008.564] [Medline: 19237720]
43. Lorig KR, Ritter PL, Dost A, Plant K, Laurent DD, McNeil I. The Expert Patients Programme online, a 1-year study of an Internet-based self-management programme for people with long-term conditions. Chronic Illn 2008 Dec;4(4):247-256. [doi: 10.1177/1742395308098886] [Medline: 19091933]
44. Lorig KR, Ritter PL, Laurent DD, Plant K. Internet-based chronic disease self-management: a randomized trial. Med Care 2006 Nov;44(11):964-971. [doi: 10.1097/01.mlr.0000233678.80203.c1] [Medline: 17063127]
45. Lorig KR, Ritter PL, Laurent DD, Plant K. The internet-based arthritis self-management program: a one-year randomized trial for patients with arthritis or fibromyalgia. Arthritis Rheum 2008 Jul 15;59(7):1009-1017 [FREE Full text] [doi: 10.1002/art.23817] [Medline: 18576310]
46. Yoon J, Zulman D, Scott JY, Maciejewski ML. Costs associated with multimorbidity among VA patients. Med Care 2014 Mar;52 Suppl 3:S31-S36. [doi: 10.1097/MLR.0000000000000061] [Medline: 24561756]
47. Roblin DW, Houston TK, Allison JJ, Joski PJ, Becker ER. Disparities in use of a personal health record in a managed care organization. J Am Med Inform Assoc 2009;16(5):683-689 [FREE Full text] [doi: 10.1197/jamia.M3169] [Medline: 19567790]
48. Song H, Cramer EM, McRoy S. Information Gathering and Technology Use Among Low-Income Minority Men at Risk for Prostate Cancer. Am J Mens Health 2014 Jun 20:-. [doi: 10.1177/1557988314539502] [Medline: 24951493]
49. Woods SS, Schwartz E, Tuepker A, Press NA, Nazi KM, Turvey CL, et al. Patient experiences with full electronic access to health records and clinical notes through the My HealtheVet Personal Health Record Pilot: qualitative study. J Med Internet Res 2013;15(3):e65 [FREE Full text] [doi: 10.2196/jmir.2356] [Medline: 23535584]

Abbreviations

CAD: coronary artery disease
CHF: congestive heart failure
COPD: chronic obstructive pulmonary disease
EHR: electronic health record
HIV: human immunodeficiency virus
ICD-9-CM: International Classification of Disease, 9th Revision, Clinical Modification
MHV: My Health*e*Vet
PHR: personal health record
PTSD: posttraumatic stress disorder
QUERI: Quality Enhancement Research Initiative
SCI: spinal cord injury
SHEP: Survey of Healthcare Experiences of Patients
TBI: traumatic brain injury
VA: US Department of Veterans Affairs

XSL•FO
RenderX

Edited by G Eysenbach; submitted 04.08.14; peer-reviewed by J Tsai, SH Tee; comments to author 04.09.14; revised version received 17.10.14; accepted 03.11.14; published 12.12.14.

Please cite as:
Shimada SL, Brandt CA, Feng H, McInnes DK, Rao SR, Rothendler JA, Haggstrom DA, Abel EA, Cioffari LS, Houston TK
Personal Health Record Reach in the Veterans Health Administration: A Cross-Sectional Analysis
J Med Internet Res 2014;16(12):e272
URL: http://www.jmir.org/2014/12/e272/
doi:10.2196/jmir.3751
PMID:25498515

©Stephanie Leah Shimada, Cynthia A Brandt, Hua Feng, D Keith McInnes, Sowmya R Rao, James A Rothendler, David A Haggstrom, Erica A Abel, Lisa S Cioffari, Thomas K Houston. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 12.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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

Consumer Participation in Quality Improvements for Chronic Disease Care: Development and Evaluation of an Interactive Patient-Centered Survey to Identify Preferred Service Initiatives

Elizabeth A Fradgley[1], BHSc(Hons); Christine L Paul[1], PhD; Jamie Bryant[1], PhD; Ian A Roos[2], PhD; Frans A Henskens[1,3], PhD; David J Paul[1,3], PhD

[1]Priority Research Centre for Health Behaviour, School of Medicine and Public Health & Hunter Medical Research Institute, University of Newcastle, Callaghan, Australia

[2]Youth Research Center, University of Melbourne, Melbourne, Australia

[3]Distributed Computing Research Group, School of Electrical Engineering and Computer Science, University of Newcastle, Callaghan, Australia

Corresponding Author:
Elizabeth A Fradgley, BHSc(Hons)
Priority Research Centre for Health Behaviour
School of Medicine and Public Health & Hunter Medical Research Institute
University of Newcastle
HMRI Building, Level 4 West
Callaghan, 2305
Australia
Phone: 61 02 4042 0642
Fax: 61 02 4042 0040
Email: elizabeth.fradgley@newcastle.edu.au

Abstract

Background: With increasing attention given to the quality of chronic disease care, a measurement approach that empowers consumers to participate in improving quality of care and enables health services to systematically introduce patient-centered initiatives is needed. A Web-based survey with complex adaptive questioning and interactive survey items would allow consumers to easily identify and prioritize detailed service initiatives.

Objective: The aim was to develop and test a Web-based survey capable of identifying and prioritizing patient-centered initiatives in chronic disease outpatient services. Testing included (1) test-retest reliability, (2) patient-perceived acceptability of the survey content and delivery mode, and (3) average completion time, completion rates, and Flesch-Kincaid reading score.

Methods: In Phase I, the Web-based Consumer Preferences Survey was developed based on a structured literature review and iterative feedback from expert groups of service providers and consumers. The touchscreen survey contained 23 general initiatives, 110 specific initiatives available through adaptive questioning, and a relative prioritization exercise. In Phase II, a pilot study was conducted within 4 outpatient clinics to evaluate the reliability properties, patient-perceived acceptability, and feasibility of the survey. Eligible participants were approached to complete the survey while waiting for an appointment or receiving intravenous therapy. The age and gender of nonconsenters was estimated to ascertain consent bias. Participants with a subsequent appointment within 14 days were asked to complete the survey for a second time.

Results: A total of 741 of 1042 individuals consented to participate (71.11% consent), 529 of 741 completed all survey content (78.9% completion), and 39 of 68 completed the test-retest component. Substantial or moderate reliability (Cohen's kappa>0.4) was reported for 16 of 20 general initiatives with observed percentage agreement ranging from 82.1%-100.0%. The majority of participants indicated the Web-based survey was easy to complete (97.9%, 531/543) and comprehensive (93.1%, 505/543). Participants also reported the interactive relative prioritization exercise was easy to complete (97.0%, 189/195) and helped them to decide which initiatives were of most importance (84.6%, 165/195). Average completion time was 8.54 minutes (SD 3.91) and the Flesch-Kincaid reading level was 6.8. Overall, 84.6% (447/529) of participants indicated a willingness to complete a similar survey again.

Conclusions: The Web-based Consumer Preferences Survey is sufficiently reliable and highly acceptable to patients. Based on completion times and reading level, this tool could be integrated in routine clinical practice and allows consumers to easily

XSL•FO
RenderX

participate in quality evaluation. Results provide a comprehensive list of patient-prioritized initiatives for patients with major chronic conditions and delivers practice-ready evidence to guide improvements in patient-centered care.

(J Med Internet Res 2014;16(12):e292) doi:10.2196/jmir.3545

KEYWORDS

ambulatory care; health care surveys; patient-centered care; consumer participation; medical oncology; chronic disease; cardiology; neurology

Introduction

Background

In the past decade, chronic diseases have become the leading cause of death worldwide and are associated with 59% of deaths and 46% of the global disease burden [1]. Prevalent chronic diseases include hypertension, diabetes mellitus, arthritis, asthma, chronic obstructive pulmonary disease, nonmelanoma cancers, and depression [2,3]. Care for chronic diseases usually requires comprehensive, personalized, and long-term services involving multidisciplinary teams. This complex care is often delivered by outpatient clinics, which are defined as services providing diagnostic or therapeutic care not requiring an overnight stay in a medical institution [4].

Within most developed countries, hospital-based outpatient clinics provide a substantial proportion of health care services and require considerable resources. For example, within Australia, hospital outpatient costs in 2011 represented approximately 61% of all health care spending [5]. The National Hospital Ambulatory Medical Care Survey reported 96.1 million outpatient department visits within the United States in 2009 alone [6]. Therefore, quality assurance initiatives targeting hospital-based outpatient services have the potential to deliver substantial benefits from both a patient perspective and a health service efficiency perspective.

A patient-centered framework is a critical component to improving chronic disease care. Patient-centered care recognizes the values, preferences, and involvement of patients and their loved ones and establishes patients as an expert information source for assessing health care quality [7]. This quality indicator has been adopted into both evaluation practice and national policies including the Australian National Health Performance Framework [8]; the United Kingdom's National Standards, Local Action, Health and Social Care Standards and Planning Framework [9]; and the Canadian Institute for Health Information's Health Indicators [10].

Appropriate measurement of patient-centered care is essential to quality evaluation practices. Patient satisfaction surveys and unmet need measures, such as the Supportive Care Need Survey [11] and the Camberwell Assessment of Needs [12], elicit participants' evaluations of outpatient care and are traditionally administered in pen-and-paper format [13]. These tools allow consumers to identify existing gaps in care and summarize perceptions of health services. For example, studies of cancer patients indicated that most were satisfied with their overall care [14], but improvement was needed regarding information, relationship, and patient care needs [15-17]. Outpatients with mental disorders reported unmet needs in psychological, relationship, and activities of daily living domains [18]. Other groups, such as patients with cardiovascular disease, reported unmet information and psychological needs [19]. Overall, such literature suggests that health care services struggle to address the needs of patients who require frequent care and have greater disease severity [20]. Results from intervention studies also suggest that attempts to translate results from needs assessment tools into practice has limited or inconsistent effects on care, outcomes, and satisfaction [13,21,22].

Practice-Ready Evidence and Consumer Engagement in Designing Health Service Initiatives

The gap in translating unmet needs to improved patient-centered care may be related to difficulties in operationalizing the results of needs assessment tools. To operationalize these data and influence practice, it is important to gather additional evidence to identify patients' preferences for changes within their health care services, strategically introduce initiatives according to patients' priorities, provide clear and feasible service-level targets for initiatives, and provide sufficient detail to design initiatives that align closely with patients' preferences and priorities. Static needs assessment tools are generally not designed to deliver such comprehensive, practice-ready, and influential data across multiple chronic conditions.

First, needs assessment tools do not enable patients to be highly specific about which unmet needs should be addressed within outpatient clinics. For example, although existing tools may facilitate patients' identification of loneliness as an unmet need, patients may not expect health professionals to provide support for this issue [13]. For unmet needs that patients do want addressed within outpatient clinics, the level of detail provided by current needs assessment tools is unlikely to be adequate. For example, parking is a frequently identified unmet need, but it is reported without specificity regarding what could be changed—spaces for clinic patients only or drop-off zones for caregivers? Without a tool capable of identifying which specific action is most likely to improve patients' experiences, health services may fail to resolve the issue. A Web-based tool with adaptive questioning would allow participants to provide information that is more detailed and personally relevant and would eliminate the time and effort required to navigate inapplicable content.

Second, an accurate method of identifying initiatives in order of priority is needed to direct limited health resources to those of greatest importance to patients. In a recent literature review of needs assessment tools frequently used in oncology care, no tool included a priority setting exercise capable of generating a comprehensive yet concise list of specific service initiatives [23]. Efforts that generate such information, such as willingness

XSL•FO
RenderX

to pay or contingent valuation, are complex tasks for participants to complete using the traditional pen-and-paper format [24]. A Web-based tool with interactive survey content can be used to efficiently examine consumers' priorities.

Third, to elicit change, tools must produce results in a form that can be readily used by health service providers and managers. Items identified by patients must be modifiable on a service level and recognized as relevant by the service providers who receive the information [25].

Finally, previous research suggests that to integrate patient-reported surveys into routine clinic practice, tools should be psychometrically robust, acceptable to patients with structured and comprehensive content, and feasible to administer in health care settings as measured by completion times and ease of administration [13]. Web-based survey software can be used to ensure these criteria are met. For example, research indicates use of this technology allows for improved readability and comprehension with simplified question formats, convenient data entry, reduction of missing data, complete timing statistics recorded by the software, and reduced administration times as compared to pen-and-paper versions [13,26,27].

Need for a Comprehensive Tool to Inform Health Service Initiatives Based on Consumers' Preferences

This study reports the development and evaluation process for an interactive Web-based tool capable of providing practice-ready, influential information suitable for designing patient-centered service initiatives for chronic disease care. This information-generating tool, the Consumer Preferences Survey, includes a set of general initiatives. Using adaptive questioning and interactive survey content, the survey also contains a comprehensive list of initiatives that are more detailed and a priority setting exercise. This will offer an alternative and efficient data collection method for identifying and introducing strategic changes to outpatient services.

This study aimed to:

1. Systematically construct a tool that (1) includes a comprehensive set of patient-centered initiatives that can be introduced at a service level, (2) allows participants to easily generate a personalized list of initiatives that would improve their experiences as an outpatient, and (3) generates practice-ready and actionable evidence in the form of a list of patient-prioritized initiatives (Phase 1).
2. Establish the following in relation to the this tool: (1) test-retest reliability, (2) patient-perceived acceptability of the survey content and delivery mode, and (3) average completion time, completion rates, and Flesch-Kincaid reading score (Phase 2).

Methods

Phase 1: Systematic Development and Pretesting of the Consumer Preferences Survey

Structured Literature and Stakeholder Review

Given the extensive qualitative work underpinning measures of need and satisfaction with patient-centered care [23,28], a literature-based approach was used to generate a comprehensive pool of item content (overarching domains and health service initiatives), preference eliciting techniques (item stems and response scales), and possible prioritization exercises.

A total of 336 articles were reviewed for item content and techniques. A total of 179 unique items and 6 unique domains were identified. Approximately 5 unique item stems were identified which incorporated concepts such as satisfaction with care, impact or value of an initiative, and perceived importance of an initiative. Four relative prioritization exercises were developed and included: ranking processes, modified willingness-to-pay questions, visual apportioning of a pie chart to respective health service initiatives, and a visual analog scale in which participants were asked to place initiatives according to importance.

Items and techniques generated by the structured literature review were circulated to 2 expert committees using an iterative process. The first committee included chronic disease physicians, health service managers from hospital-based specialist services, community-based chronic disease experts, and health behavior researchers including an academic biostatistician and a health economist (n=20). The second committee included consumer advocates and health service users within cancer, neurology, and cardiology fields (n=27).

Final Survey Content

After 2 rounds of feedback from each expert group, a total of 23 general service initiatives were generated from the item pool (Table 1). These initiatives were organized as 4 steps in the process of care: (1) scheduling an appointment, (2) arriving at an appointment, (3) during a clinical appointment, and (4) managing a chronic illness at home (Figure 1). Both expert committees preferred this approach. By allowing participants to sequentially consider each way in which care was experienced and could be improved, recall bias and cognitive burden may be reduced. However, these areas are not considered latent variables or constructs.

Participants selected initiatives that would greatly improve their experience within the outpatient clinic from lists presented on the touchscreen computer. This is considered a dichotomous response scale. Selected initiatives were recorded as 1 and unselected initiatives were coded as zero. If a general initiative was selected, the survey displayed a subsequent list of specific health service initiatives using adaptive questioning: "On the last screen, you indicated that [general health service initiative] could improve your experience (Figures 2 and 3. What specifically could the clinic change to help you? [list of possible specific health services initiatives displayed]." A total 110 specific health service initiatives were available, including having emergency appointments available within a week (making an appointment), being informed of estimated wait times on arrival (arriving at an appointment), an action plan is created to address your concerns (during an appointment), and knowing which symptoms require emergency attention (self-management at home). Only those participants who selected all 23 broad health service initiatives would view all 110 detailed health service initiatives. Complete survey content is available in Multimedia Appendix 1.

XSL•FO
RenderX

Once the 4 steps were completed, participants who selected more than 5 general initiatives were presented with an autopopulated list of their previous selections and were asked to choose the 5 initiatives that were of greatest importance to them. These participants and those individuals who selected at least 2 but less than 5 initiatives were directed to a modified relative prioritization exercise. Participants were asked to allocate 100 points across their desired initiatives to indicate the relative perceived importance of each (Figures 4 and 5).

Table 1. Consumer Preferences Survey content by area of care, including number of possible specific initiatives.

Area of care and general initiatives	Specific initiatives (n)
1. Making an appointment	
Schedule convenient times	7
Easy contact with clinic staff	2
Transportation to appointment	3
2. Arriving at an appointment	
Car parking	7
Comfortable wait rooms	10
Reduced time in wait rooms	3
Having friends or family with you	—
3. During clinical appointment	
Provide more information	4
Ensure all your concerns are addressed	4
Involve you more in treatment decisions	3
Keep you up-to-date on the progress of your treatment and condition	3
Ensure good interactions and relationships with all clinic staff	5
Provide comfortable and pleasant treatment rooms	10
Provide good quality hospital catering	—
Better coordination of your care and information	7
Minimize pain or discomfort when you receive treatment	4
4. Managing at home	
Help with physical symptoms	6
Help with emotional health or relationships	4
Help with daily activities and healthy lifestyles	6
Help with employment, finances, or insurance	5
Information on your condition and treatment	8
Support and involvement of family and friends	5
Knowledge on how to handle a medical emergency	4

XSL•FO
RenderX

Figure 1. Screenshot of the Consumer Preferences Survey, introduction and instruction screen.

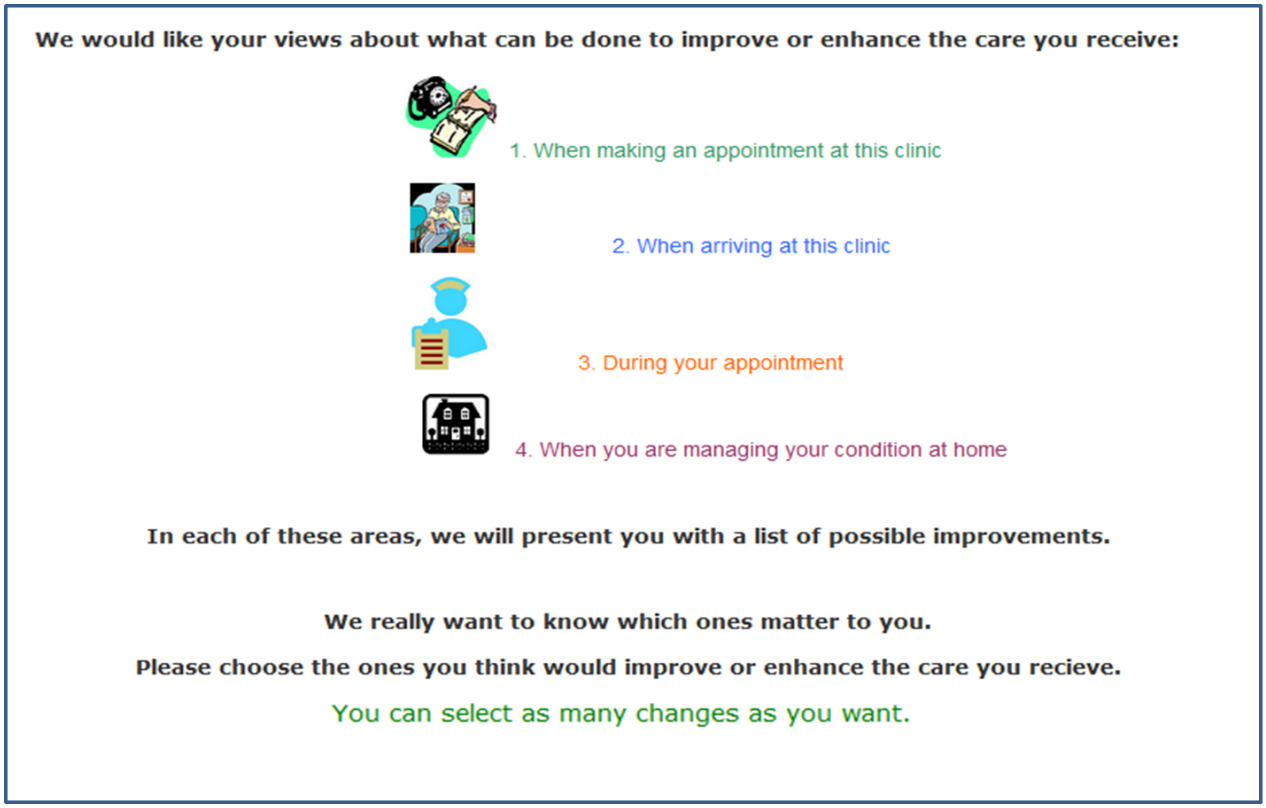
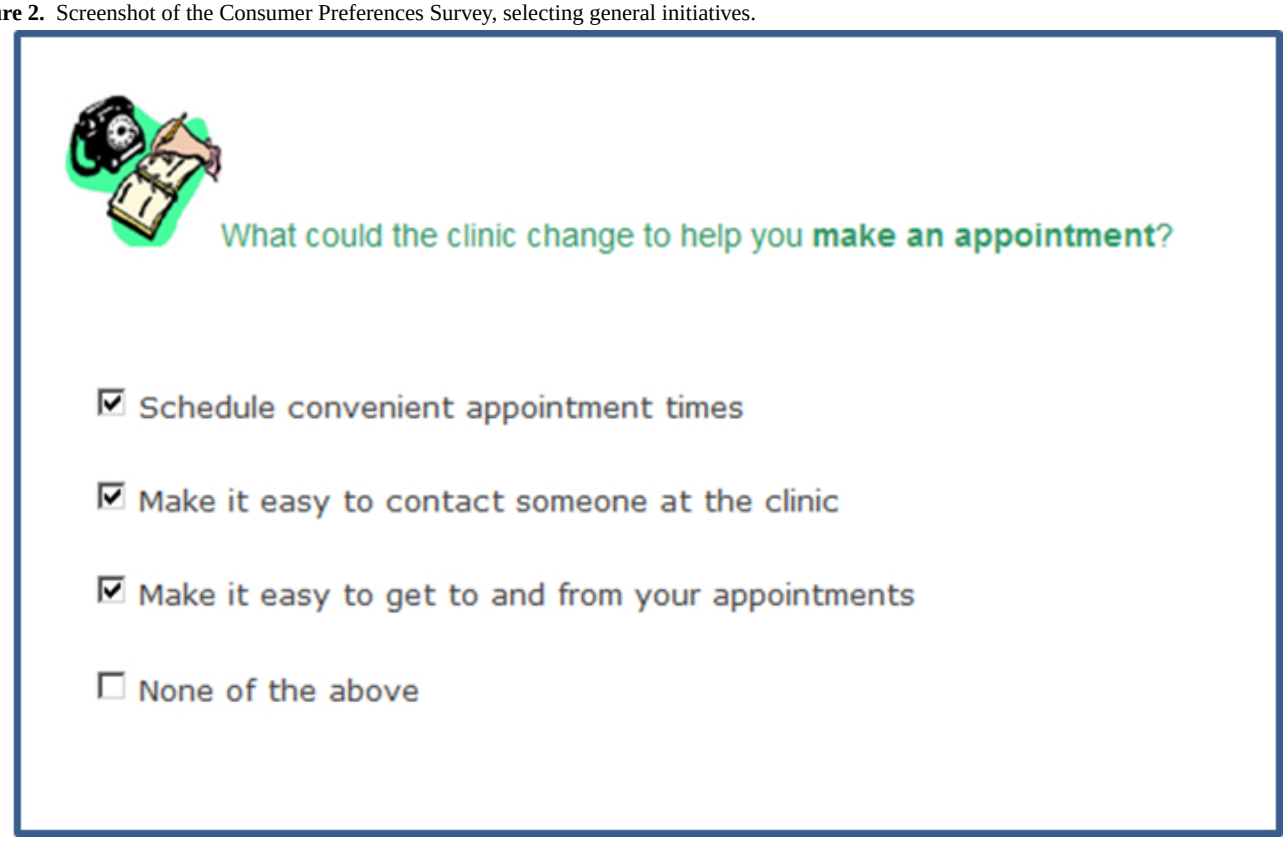
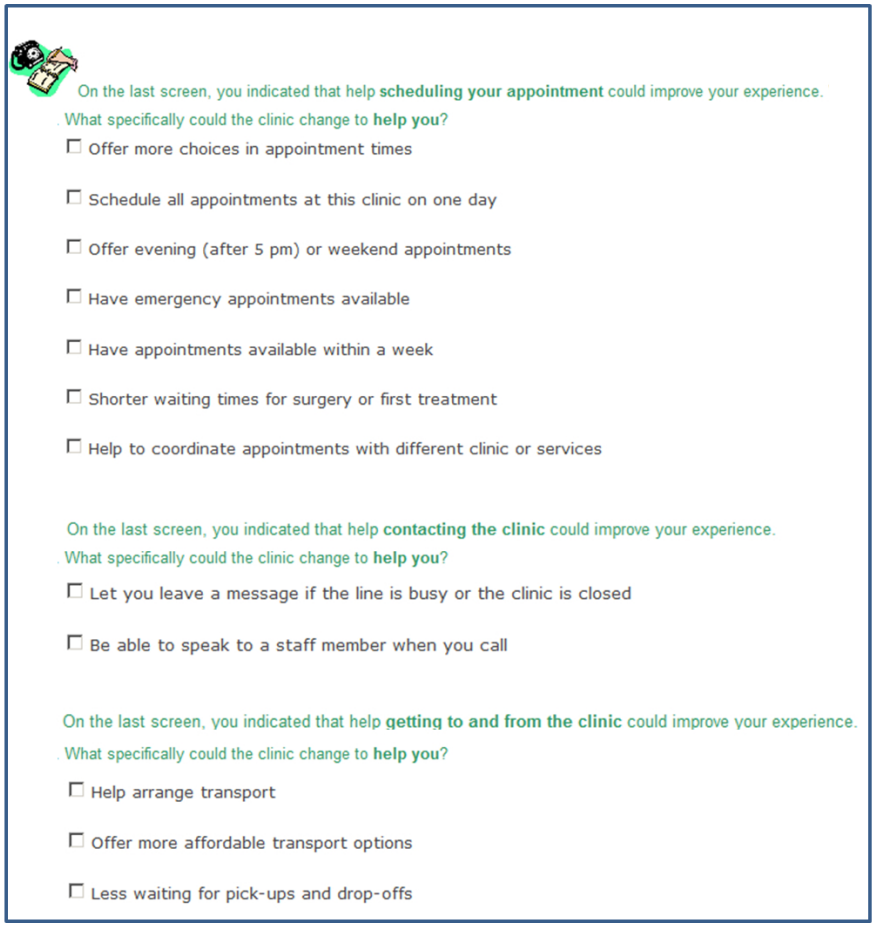


Figure 2. Screenshot of the Consumer Preferences Survey, selecting general initiatives.



XSL•FO
RenderX

Figure 3. Screenshot of the Consumer Preferences Survey, selecting specific initiatives.

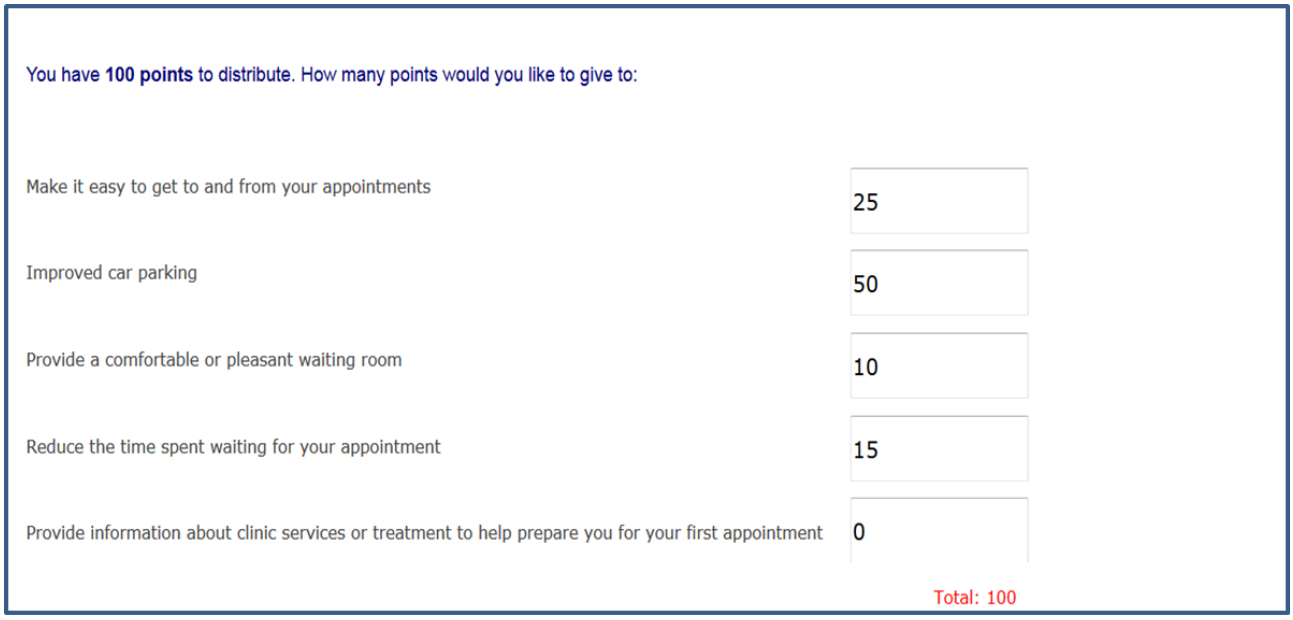


XSL•FO
RenderX

Figure 4. Screenshot of Consumer Preferences Survey, instructions for relative prioritization exercise.



Figure 5. Screenshot of Consumer Preferences Survey, relative prioritization exercise.



XSL•FO
RenderX

Final Web-Based Format

To facilitate adaptive questioning and the branching patterns required to navigate the survey, a novel software program using touchscreen technology was developed in collaboration with health behavior researchers and information technology experts [29]. To confirm the technical functionality and usability of the survey software, a total of 75 participants pretested the final format without error.

To commence the survey, a research assistant first registered the user with a unique username before handing the touchscreen device to the participant. This registration step allowed participants to pause and restart the survey without losing previously entered information by re-entering their username. The unique username was stored with each result set, and removed before analysis, to allow detection of nonunique participants. If a duplicate was discovered, the entry with complete data only was used for analysis.

Once a username was created, participants were able to progress through the survey using a clearly presented "Next" button located at the bottom of every survey page. Participants were also able to navigate to previous responses using the "Back" button. All participants received 4 instruction screens and 4 screens listing the 23 general initiatives. In the unlikely scenario a participant selected all general initiatives, based on adaptive questioning they would receive an additional 7 screens listing 110 specific initiatives. The prioritization exercise included 2 instruction screens and 2 exercise screens. The maximum number of survey items presented on a screen was 4 and participants may have been required to page scroll to view all items.

Phase 2: Test-Retest Reliability and Patient Acceptability of the Consumer Preferences Survey

Clinic Settings

High-volume tertiary outpatient medical oncology, cardiology, and neurology clinics were included to ensure the pilot sample reflected a range of prevalent chronic illnesses. Pilot sites included a public tertiary outpatient clinic specializing in both cardiology and neurology care, a public tertiary outpatient clinic specializing in oncology care, and a private tertiary outpatient clinic specializing in intravenous chemotherapy only.

Participant Eligibility

Eligible participants were able to read English, 18 years of age or older, and had attended the clinic at least once prior to recruitment. A subsample of participants completed the survey again within 14 days to assess test-retest reliability. Eligibility for the test-retest component of the study required written consent and a second appointment scheduled within 10-14 days. Given this narrow timeframe, only medical oncology patients with an ongoing treatment schedule were approached to participate in the test-retest component of the study.

Recruitment and Survey Administration

Trained research assistants approached patients in the outpatient clinic waiting rooms or intravenous chemotherapy treatment spaces. Eligible participants were invited to complete the survey at the time of recruitment only and individuals were not provided the website address to access the survey outside of the clinic setting. The survey was voluntary, not advertised, and no incentives to participate were offered. The age and gender of nonconsenters was estimated to ascertain consent bias.

Measures

The touchscreen survey consisted of the Consumer Preferences Survey and the following:

1. Demographic information: age, gender, marital status, education, private health insurance, concessional card, Aboriginal or Torres Strait Islander origin, and appointment frequency within the past 3 months were collected. Participants also reported the reason for attending the clinic with response options of a routine exam for a diagnosed condition, discussion of symptoms for a diagnosed or nondiagnosed condition, or to receive tests or treatments.
2. Acceptability items: a total of 6 questions assessed the acceptability of the Consumer Preferences Survey: (1) Do you think the directions given for filling out the survey were adequate?, (2) Overall, how would you rate the length of the survey? with response options "it was too short," "it was just right," or "it was too long," (3) Did the survey miss any changes that could improve your experience in this outpatient clinic?, (4) Did you find filling in the survey confusing or difficult?, (5) Would you be willing to complete a similar survey in the future?, and (6) Do you believe the survey will provide an accurate summary of initiatives which could improve your experience within the outpatient clinic? Only those participants who reported difficulty completing the survey were asked additional questions assessing the ease of navigation, layout or functioning of iPad screens, adequacy of directions, and whether some changes would be helpful. Participants who selected at least 2 general initiatives and were instructed to complete the relative prioritization exercise received 3 additional questions, including (1) Do you think the directions given for this exercise were adequate?, (2) Did this exercise help you to decide which changes to the clinic are most important to you?, and (3) Did you find this exercise difficult?

Data Analysis

To examine test-retest reliability, nonweighted Cohen's kappa coefficients and percent agreements were calculated to report differences between responses at participants' first completion of the survey and second completion of the survey. Items with a kappa value equal to or greater than .60 were considered to have substantial test-retest reliability [30]. Those items reporting a kappa value from .40 to .59 were considered to have moderate test-retest reliability.

Acceptability items were examined using proportions and 95% confidence intervals. Differences in estimated age, gender, or clinic characteristics of consenters or nonconsenters were examined using chi-square statistics. A P value of <.05 was considered statistically significant.

Ease of integration of the Consumer Preferences Survey was assessed by examining Flesch-Kincaid reading level [31],

XSL•FO
RenderX

average time to complete, and survey completion rates. The survey software recorded timing statistics and survey completion rates. The average time to complete, including standard deviations, each portion of the survey is reported.

Institutional Review Board Approval and Data Protection

Ethics approval was provided by Hunter New England Human Research Ethics Committee and the University of Newcastle Human Research Ethics Committee. Consent was implied if an individual chose to begin the survey. All personal information was immediately uploaded via an encrypted channel and stored on secure university servers with password-protected access granted to study researchers only.

Results

Summary

A total of 1042 chronic disease outpatients were approached to participate over a 10-month period from March to December 2013. A total of 741 individuals agreed to participate—a 71.11% consent rate (Table 2). Of the 301 individuals who declined to participate, clinic site was documented and age and gender estimated for 291 individuals (96.7%). Of the 741 consenting participants, age, gender, and clinic sites were recorded for 674 individuals (91.0%). There were no significant differences between consenters and nonconsenters by gender (*P*=.85). Age category was significantly related to consent (*P*=.007). Consent rates were also significantly higher within the privately funded intravenous chemotherapy clinic compared with both the publically funded oncology clinic and publically funded cardiology and neurology clinic (*P*=.001).

A total of 143 of 184 medical oncology participants (consent rate 78.1%) were willing to participate in the test-retest component. Only 68 of these 143 had a scheduled appointment within 14 days (48.9% eligibility). Due to rescheduled appointments and participants' physical well-being at the second appointment, 39 participants were included in the final test-retest sample.

Table 2. Demographic characteristics by consent status for Consumer Preferences Survey pilot (N=965).

Demographic characteristic	Nonconsenters, n (%) (n=291)	Consenters, n (%) (n=674)	χ^2 (*df*)	*P*
Male (n=438)	134 (46.2)	304 (45.1)	0.04 (1)	.85
Clinic site			14.7 (2)	.001
Public oncology (n=476)	148 (31.1)	328 (68.9)		
Public cardiology and neurology (n=415)	135 (32.5)	280 (67.5)		
Private oncology (n=74)	8 (10.8)	66 (89.2)		
Age category			14.1 (4)	.007
18-25 (n=38)	5 (2.7)	33 (4.6)		
26-40 (n=155)	49 (16.8)	106 (16.1)		
41-55 (n=262)	80 (27.9)	182 (25.4)		
56-70 (n=354)	95 (31.9)	259 (35.8)		
≥71 (n=156)	62 (20.8)	94 (18.2)		

Sample Demographic Characteristics

A total sample of 674 participants completed the demographic module and included 394 medical oncology patients (58.5%) and 280 (41.5%) cardiology or neurology patients (Table 3). Females were slightly overrepresented (54.9%, 370/674) in the sample and the average age was approximately 59.7 years (SD 15.5 years). Participants were most likely to have attained a high school equivalent of year 10 or lower (49.2%, 332/674) and to be married or living with a partner (66.3%, 447/674). The most common reported reasons for attending the clinic were related to a diagnosed condition, with 41.9% (282/674) of participants attending for a routine exam and 30.5% (206/674) attending to receive tests or treatment.

XSL•FO
RenderX

Table 3. Sample demographic characteristics of Consumer Preferences Survey pilot test (N=674).

Sample characteristics	Participants
Age (years), mean (SD)	59.7 (15.5)
Male, n (%)	304 (45.1)
Highest level of education attained, n (%)	
High school equivalent of year 10 or lower	332 (49.2)
High school completion	93 (13.8)
Diploma or trade certificate	140 (20.8)
Bachelor's degree	63 (9.4)
Not specified	46 (6.8)
Marital status, n (%)	
Married or living with partner	447 (66.3)
Single	84 (12.5)
Widowed	84 (12.5)
Not specified	59 (8.7)
Aboriginal and/or Torres Strait Islander origin, n (%)	28 (4.2)
No private insurance coverage, n (%)	361 (53.6)
Concessional card, n (%)	448 (66.5)
Chronic condition group, n (%)	
Cardiology or neurology	280 (41.5)
Medical oncology	394 (58.5)
Medical oncology private facility, n (%)	66 (9.8)
Reason for attending, n (%)	
To discuss symptoms, treatments or tests for diagnosed condition	121 (17.9)
To discuss symptoms or tests for undiagnosed condition	49 (7.3)
To receive tests or treatments for diagnosed condition	206 (30.5)
For a routine exam for a diagnosed condition	282 (41.9)
Not specified	16 (2.4)
Number of appointments in last 3 months, n (%)	
1 in last 6 months	328 (48.7)
2-3	145 (21.5)
4-5	92 (13.6)
6	31 (4.6)
≥7	73 (10.8)
Not specified	3 (0.8)

Reliability Statistics

A total of 39 oncology patients participated in the test-retest component (Table 4). Substantial test-retest reliability was reported for 9 general initiative items (Cohen's kappa>.6) and moderate test-retest reliability was reported for 7 general initiatives (Cohen's kappa=.40-.59). Four initiatives reported a value below a .4 threshold, indicating poor reliability. However, observed agreement for these items ranged from 94.9%-97.4%. There were an insufficient number of observations to calculate a test statistic for 3 initiatives. Across all initiatives, the average observed agreement was 93.7% with moderate test-retest reliability (Cohen's kappa=.53).

XSL•FO
RenderX

Table 4. Cohen's kappa values and observed percentage agreement for general initiatives (n=39).

General initiatives selected by area of care	Observed agreement (%)	Cohen's κ (95% CI)
1. Area of care: making an appointment		
Provide more convenient appointment times	89.7	.44 (–.01, .09)
Make it easier to contact the clinic	94.9	.64 (.18, 1.00)
Help to arrange transport to and from the clinic	94.9	.47 (.15, 1.00)
None selected	89.7	.69 (.40, .97)
Total number selected	87.2	.62 (.46, .68)
2. Area of care: arriving at an appointment		
Improve car parking	89.7	.79 (.59, .98)
Provide a comfortable and pleasant waiting room[a]	—	—
Reduce waiting times	94.9	.47 (–.15, 1.00)
Ensure family and friends are comfortable within waiting rooms	97.4	.66 (.03, 1.00)
None selected	82.1	.64 (.40, .88)
Total number selected	84.6	.71 (.52, .95)
3. Area of care: arriving at an appointment		
Provide more information about treatment and condition	97.4	0
Ensure your concerns are discussed with health care professionals	97.4	.79 (.38, 1.00)
Involve you in treatment decisions	100.0	1.00 (1.00, 1.00)
Keep you up-to-date on your treatment and condition progress	92.3	.53 (.06, .99)
Ensure good interactions with all clinic staff	97.4	0
Provide a comfortable and pleasant treatment room[a]	—	—
Provide good hospital catering	89.7	.55 (.18, .92)
Better coordination of your care	97.4	.66 (.03, 1.00)
Minimize pain or discomfort during treatment[a]	-	—
None selected	84.6	.60 (.32, .88)
Total number selected	79.5	.50 (.33, .80)
4. Area of care: managing your condition at home		
Access to help or information to manage physical symptoms	97.4	.66 (.03, 1.00)
Access to help or information to manage emotional symptoms	94.9	–.03 (–.09, .04)
Access to help in order to maintain activities of daily living	94.9	–.03 (–.09, .04)
Access to help or information relating to finance, work, insurance	97.4	.84 (.54, 1.00)
Access to information to review at home	94.9	.48 (–.12, 1.00)
Access to help or information for family support	94.9	.48 (–.12, 1.00)
Information on how to manage medical emergencies	94.9	.64 (.19, 1.00)
None selected	89.7	.72 (.46, .97)
Total number selected	84.6	.60 (.29, .63)

[a] Insufficient number of observations to calculate a test statistic.

Acceptability Statistics

A total of 543 of 674 individuals (80.6%) completed the acceptability items related to selecting and navigating general initiatives and 529 (78.9%) completed items related to the perceived value of survey results (Table 5). This noncompletion rate was observed for those participants who were called into their appointment before completing the survey. Of the 543 participants, most found the Consumer Preferences Survey easy to complete (97.9%, 531/543), comprehensive (93.1%, 506/543), an appropriate length (95.5%, 519/543), and thought adequate directions were provided (98.3%, 534/543). Of the 195 participants who received the relative prioritization exercise,

XSL•FO
RenderX

the majority indicated it was easy to complete (97.0%, 189/195) and that directions were clear (94.6%, 184/195).

A minority of participants indicated they were unsure if the results were an accurate summary of the initiatives desired (17.4%, 92/529) or were unsure if they would be willing to complete a similar survey in future (9.9%, 52/529). However, the majority of participants (80.7%, 427/529) believed the survey results were an accurate summary of initiatives that could improve their experience within the outpatient clinic, whereas 84.6% (165/195) reported the relative prioritization exercise helped them to decide which general initiatives were of greatest importance. Overall, 84.6% (448/529) of participants were willing to complete a similar survey in the future. The Flesch-Kincaid reading level was grade 6.8, indicating that those who completed 7 years of formal education would be able to easily comprehend the survey content.

Table 5. Reported acceptability of the Consumer Preferences Survey, including relative prioritization exercise.

Acceptability	Participants, % (95% CI)
Selecting and navigating general initiatives (n=543)	
The directions provided were adequate	98.3 (97.2-99.6)
The length of the survey was appropriate	95.5 (93.5-97.5)
The survey was comprehensive of all initiatives	93.1 (90.7-95.6)
The survey was clear and easy to complete	97.9 (96.4-99.3)
Completing the relative prioritization exercise[a] (n=195)	
The directions for the points exercise was adequate	94.6 (92.2-97.0)
The point exercise helped to decide what was important	84.6 (80.4-88.8)
The point exercise was clear and easy to complete	97.0 (95.2-98.8)
Overall value of survey (n=529)	
The survey is an accurate summary of the initiatives desired	80.7 (77.0-84.5)
Willing to complete a similar survey in the future	84.6 (81.0-88.2)

[a] Completed by only those participants with 2 or more general initiatives selected.

Completion Rates and Timing Statistics

Approximately 78.4% (529/674) of participants completed the Consumer Preferences Survey and all acceptability questions. Completion was significantly related to clinic site (data not shown; $P<.001$), with completion rates significantly higher within the privately funded intravenous chemotherapy clinic (98.5%, 65/66) compared to the publically funded oncology clinic (82.3%, 325/394). Both oncology clinics reported significantly higher completion than publically funded cardiology and neurology clinic (69.3%, 194/280).

Approximately 5 minutes (mean 5.02, SD 3.07) was required to navigate and select initiatives and to complete the relative prioritization exercise. The total time to complete all pilot survey content, excluding acceptability questions, was approximately 8.54 (SD 3.91) minutes.

Discussion

Principal Results

The development of the Consumer Preferences Survey was successful in providing a novel tool capable of generating a personalized and concise list of health service initiatives relevant to patients' experiences of outpatient care, identifying a comprehensive set of targets that are modifiable on a service level, and generating a list of prioritized initiatives to ensure service-level change is introduced strategically. The interactive survey software also allows participants to select up to 110 specific initiatives and indicate the relative importance of chosen initiatives in improving their care experience.

Results from our pilot study suggest the tool is sufficiently reliable and acceptable to patients. The test-retest reliability of each general initiative was moderate to substantial and observed percentage agreement was above 80%, indicating that this tool provides a stable summary of patients' preferences for health service change. Participants reported the Consumer Preferences Survey was easy to complete, comprehensive, and of an appropriate length. Based on average completion times and reading level, this tool can also be integrated into routine clinic practice and allows consumers to quickly participate in a quality evaluation exercise. Time to complete is approximately 9 minutes and is comparable to, or shorter than, other patient-report tools, such as the Cancer Care Monitor (12 minutes), Supportive Care Needs Survey (15-20 minutes), and Camberwell Assessment of Need (16 minutes) [12,13]. The Flesch-Kincaid reading level of the survey was assessed at 6.8. This requires participants to have completed a level of formal education well below the level of education recommended by the Australian National Health Medical Research Council for presenting information to health consumers [32].

Electronic touchscreen surveys are becoming a popular mode of survey administration within health research [33]. The Consumer Preferences Survey uses innovative Web-based software capable of complex adaptive questioning and interactive item types. The branching patterns allow participants to easily navigate through all general initiatives and only receive

XSL•FO
RenderX

subsequent questions focusing on specific initiatives when appropriate. The survey content, particularly the way in which participants receive questions and the relative prioritization exercise, is a novel approach to summarizing and prioritizing patients' perceptions of the quality of care. As such, it is promising that approximately 85% of individuals indicated they would be willing to complete a similar survey in the future and only 2% believed the survey did not provide an accurate summary of desired health service initiatives. The relative prioritization exercise was perceived to be a helpful and easy exercise that could be completed in a relatively short amount of time. Similar Web-based exercises have been successfully used to explore consumers' research priorities and decision preferences [24,34]. With limited health care resources available, simplified willingness-to-pay exercises may be an appropriate data collection approach to strategically determine funding priorities according to consumers' preferences.

Value and Application of the Consumer Preferences Survey

The involvement of consumers in shaping health policy and services is widely recognized as important for promoting patient-centered care in chronic diseases. Firstly, it is seen as an ethical and democratic right [35]. This can be an empowering experience for consumers who have been traditionally passive receivers of care with little opportunity to discuss their experiences. Secondly, consumers also offer a different but equally important perspective on the quality of health care than those of health service providers, researchers, and policy makers [35]. As research funding organizations and government health policies continue to mandate consumer involvement, a process to systematically collect and measure consumers' perspectives of care is needed.

The Consumer Preferences Survey allows consumers to participate in a quality evaluation exercise and provides valuable information on how health services can be restructured. This is an information-generating tool and can be used to determine consumers' preferences and priorities for health service initiatives. The data collection process is systematic, meaningful to consumers and health services, and sufficiently detailed and concise to translate into meaningful patient-centered health initiatives. The touchscreen survey covers a wide range of issues relatively quickly to minimize participant burden and maximize the feasibility of administering the tool in a range of health care settings.

Limitations of the Consumer Preference Survey

Unlike tools such as the Patient Generated Index [36], the Consumer Preferences Survey does not allow participants to nominate other aspects of their care they would like changed. However, these tools are not amenable to touchscreen mode of administration and cannot incorporate benefits such as an interactive action-prioritization exercise. Furthermore, a format allowing participants to generate new (and potentially highly personalized) items introduces challenges in relation to the prioritization of initiatives across chronic condition groups and users. As part of the evaluation process, the pilot test allowed participants to suggest additional initiatives that had not been covered by the survey. Only 5% of individuals suggested an additional initiative. Suggested initiatives were often specific initiatives that had been eliminated by expert review because they were deemed nonmodifiable or relevant to only a very small portion of service users.

The Consumer Preferences Survey has not been tested for validity following some traditional psychometric methods, such as factor analysis, given the lack of common denominator for items due to use of a sophisticated branching pattern. Furthermore, the organization structure of the survey into 4 distinct chronological areas of care eliminates the possibility of item randomization. These areas of care serve only as an organization structure and do not infer latent variables or construct dimensions. However, given that the Consumer Preference Survey was not designed to measure a psychological construct or diagnose a disease state, but rather as a tool for identifying and prioritizing subjective changes to hospital-based outpatient care, reliability is likely to be the most appropriate and important psychometric characteristic to establish. To further establish the credibility of survey results, additional research replicating the reliability portion of this pilot study with a greater sample size is required.

Although this pilot study included a range of settings and a large number of participants, there are several methodological limitations that may introduce bias. Firstly, age and gender were not recorded for all consenting individuals (9% missing data). As described, participants were recruited in clinic waiting rooms before their appointment or in treatment rooms while receiving intravenous therapy. All participants recruited within the privately funded clinic completed the survey while receiving intravenous treatment and did not experience time constraints. Missing data are observed for those participants called into their appointment before survey completion; this predominately occurred within the publically funded clinics. This difference in recruitment location explains the significant difference between clinic site and completion rates. There are a number of benefits that justify applying an active recruitment method within the health service settings. Firstly, participants can use the touchscreen device and survey software specifically designed for this study instead of more laborious methods, such as a postal survey or arranging a telephone interview. Secondly, participation rates are much higher with face-to-face interaction and may mitigate any response bias [37]. This can result in a more equitable representation of patients' preferred initiatives. Although it could be argued that this approach introduces social desirability bias, this may be mitigated by the touchscreen technology which prior research suggests is a very private and acceptable data collection method [38].

Within the pilot, age was significantly related to consent. Individuals aged 71 years or older were more likely to decline participation than any other age group. There is evidence suggesting that age is generally not a barrier to using touchscreen technology within ambulatory settings [33]; within the study, the second oldest age group (50-69 years) reported a slightly higher than average consent rate (73%). This may suggest this result is isolated or due to an additional explanatory variable, such as increasing illness severity within the older age group. However, this cannot be confirmed within the existing dataset and requires further evaluation.

XSL•FO
RenderX

Conclusions

Results from the Consumer Preferences Survey can be used to guide patient-centered care initiatives within health services and will provide a list of patient-prioritized targets across several chronic conditions. This will offer an alternative and reliable method to introduce strategic initiatives to chronic disease outpatient services with the objective of empowering consumers to participate in quality improvement activities.

Acknowledgments

Dr Jamie Bryant is supported by an Australian Research Council Post-Doctoral Industry Fellowship. Elizabeth Fradgley is supported by a Strategic Research Partnership Grant from the Cancer Council NSW to the Newcastle Cancer Control Collaborative. A/Prof Christine Paul was supported by an HMRI Fellowship and an NHMRC Career Development Fellowship. Infrastructure funding is provided by the Hunter Medical Research Institute and this project is a component of an Australian Research Council Team Linkage Project. Authors acknowledge the valuable contribution of individuals at Hunter Medical Research Institute (Team HMRI), CReDITSS (Dr. Chris Oldmeadow), Cancer Council New South Wales, staff within Hunter New England Health and HealthScope, and patients who kindly provided feedback on survey development.

Authors' Contributions

EF, JB, and CP participated in study design, management, and data analysis. IR facilitated content development and stakeholder review. FH and DP led Web-based software development. All authors participated in development of the Consumer Preference Survey. All authors drafted and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Screenshots of the Consumer Preferences Survey.

[PDF File (Adobe PDF File), 550KB - jmir_v16i12e292_app1.pdf]

References

1. Coleman K, Austin BT, Brach C, Wagner EH. Evidence on the Chronic Care Model in the new millennium. Health Aff (Millwood) 2009;28(1):75-85 [FREE Full text] [doi: 10.1377/hlthaff.28.1.75] [Medline: 19124857]
2. Schoen C, Osborn R, How SK, Doty MM, Peugh J. In chronic condition: experiences of patients with complex health care needs, in eight countries, 2008. Health Aff (Millwood) 2009;28(1):w1-16 [FREE Full text] [doi: 10.1377/hlthaff.28.1.w1] [Medline: 19008253]
3. Tritter JQ, Calnan M. Cancer as a chronic illness? Reconsidering categorization and exploring experience. Eur J Cancer Care (Engl) 2002 Sep;11(3):161-165. [Medline: 12296831]
4. Shi L, Singh DA. Delivering Health Care in America: A Systems Approach. Sudbury, MA: Jones and Bartlett; 2008.
5. Australian Institute of Health and Wellness. Health Services Series. Canberra: Australian Institute of Health and Wellness; 2012 Apr. Australian hospital statistics: 2010-11 (Cat No. HSE 117) URL: http://www.aihw.gov.au/publication-detail/?id=10737421633 [accessed 2014-05-15] [WebCite Cache ID 6PZoEfzgn]
6. Barnett ML, Song Z, Landon BE. Trends in physician referrals in the United States, 1999-2009. Arch Intern Med 2012 Jan 23;172(2):163-170 [FREE Full text] [doi: 10.1001/archinternmed.2011.722] [Medline: 22271124]
7. Committee on Quality of Health Care in America, Institute of Medicine. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: National Academy Press; 2001.
8. National Health Performance Committee. Australian Institute of Health and Welfare. 2001. National health performance framework report (Cat. no. AIHW 7112) URL: http://www.aihw.gov.au/publication-detail/?id=6442467275 [accessed 2014-05-15] [WebCite Cache ID 6PZoPZRQl]
9. Department of Health. National Standards, Local Action. Health and Social Care Standards and Planning Framework. 2005. URL: http://webarchive.nationalarchives.gov.uk/20130107105354/http://dh.gov.uk/en/publicationsandstatistics/publications/publicationspolicyandguidance/dh_4086057 [accessed 2014-05-15] [WebCite Cache ID 6PZocHV2Y]
10. Canadian Institute for Health Information. Pan-Canadian Primary Health Care Indicators. Ottawa, ON: Canadian Institute for Health Information; 2006. URL: http://www.cihi.ca/CIHI-ext-portal/pdf/internet/PDF_PHC_INDI_REPORT1VOL1FIN_EN [accessed 2014-12-01] [WebCite Cache ID 6UU7ul3S4]
11. Bonevski B, Sanson-Fisher R, Girgis A, Burton L, Cook P, Boyes A. Evaluation of an instrument to assess the needs of patients with cancer. Supportive Care Review Group. Cancer 2000 Jan 1;88(1):217-225. [Medline: 10618626]

XSL•FO
RenderX

12. Phelan M, Slade M, Thornicroft G, Dunn G, Holloway F, Wykes T, et al. The Camberwell Assessment of Need: the validity and reliability of an instrument to assess the needs of people with severe mental illness. Br J Psychiatry 1995 Nov;167(5):589-595. [Medline: 8564313]
13. Richardson A, Sitzia J, Brown V, Medina J, Richardson A. Patients' needs assessment tools in cancer care: principles and practice. London, UK: King's College London; 2005. URL: http://www.kcl.ac.uk/teares/nmvc/external/docs/ar-patients-needs-assessment-final.pdf [accessed 2014-05-15] [WebCite Cache ID 6PZpCn6dK]
14. Heading G, Mallock N, Sinclair S, Bishop J. Cancer Patient Satisfaction Survey. Sydney, Australia: Cancer Institute of New South Wales; 2009. URL: http://www0.health.nsw.gov.au/pubs/2009/pdf/2009_patient_survey_report.PDF [accessed 2014-05-15] [WebCite Cache ID 6PZpI1HXF]
15. Harrison JD, Young JM, Price MA, Butow PN, Solomon MJ. What are the unmet supportive care needs of people with cancer? A systematic review. Support Care Cancer 2009 Aug;17(8):1117-1128. [doi: 10.1007/s00520-009-0615-5] [Medline: 19319577]
16. Morrison V, Henderson BJ, Zinovieff F, Davies G, Cartmell R, Hall A, et al. Common, important, and unmet needs of cancer outpatients. Eur J Oncol Nurs 2012 Apr;16(2):115-123. [doi: 10.1016/j.ejon.2011.04.004] [Medline: 21555246]
17. Sanson-Fisher R, Girgis A, Boyes A, Bonevski B, Burton L, Cook P. The unmet supportive care needs of patients with cancer. Supportive Care Review Group. Cancer 2000 Jan 1;88(1):226-237. [Medline: 10618627]
18. Goossens PJ, Knoppert-van der Klein EA, Kroon H, van Achterberg T. Self-reported care needs of outpatients with a bipolar disorder in the Netherlands. J Psychiatr Ment Health Nurs 2007 Sep;14(6):549-557. [doi: 10.1111/j.1365-2850.2007.01127.x] [Medline: 17718727]
19. Kohlmann S, Kilbert MS, Ziegler K, Schulz KH. Supportive care needs in patients with cardiovascular disorders. Patient Educ Couns 2013 Jun;91(3):378-384. [doi: 10.1016/j.pec.2013.01.002] [Medline: 23391685]
20. McColl MA, Jarzynowska A, Shortt S. Unmet health care needs of people with disabilities: population level evidence. Disability & Society 2010 Mar;25(2):205-218. [doi: 10.1080/09687590903537406]
21. Kazis LE, Callahan LF, Meenan RF, Pincus T. Health status reports in the care of patients with rheumatoid arthritis. J Clin Epidemiol 1990;43(11):1243-1253. [Medline: 2243259]
22. Snyder CF, Dy SM, Hendricks DE, Brahmer JR, Carducci MA, Wolff AC, et al. Asking the right questions: investigating needs assessments and health-related quality-of-life questionnaires for use in oncology clinical practice. Support Care Cancer 2007 Sep;15(9):1075-1085. [doi: 10.1007/s00520-007-0223-1] [Medline: 17318591]
23. Richardson A, Medina J, Brown V, Sitzia J. Patients' needs assessment in cancer care: a review of assessment tools. Support Care Cancer 2007 Oct;15(10):1125-1144. [doi: 10.1007/s00520-006-0205-8] [Medline: 17235503]
24. Matthews D, Rocchi A, Wang EC, Gafni A. Use of an interactive tool to assess patients' willingness-to-pay. J Biomed Inform 2001 Oct;34(5):311-320 [FREE Full text] [doi: 10.1006/jbin.2002.1032] [Medline: 12123150]
25. Snyder CF, Garrett-Mayer E, Brahmer JR, Carducci MA, Pili R, Stearns V, et al. Symptoms, supportive care needs, and function in cancer patients: how are they related? Qual Life Res 2008 Jun;17(5):665-677 [FREE Full text] [doi: 10.1007/s11136-008-9331-8] [Medline: 18493865]
26. Velikova G, Wright EP, Smith AB, Cull A, Gould A, Forman D, et al. Automated collection of quality-of-life data: a comparison of paper and computer touch-screen questionnaires. J Clin Oncol 1999 Mar;17(3):998-1007. [Medline: 10071295]
27. Boyes A, Newell S, Girgis A. Rapid assessment of psychosocial well-being: are computers the way forward in a clinical setting? Qual Life Res 2002 Feb;11(1):27-35. [Medline: 12003053]
28. Hudon C, Fortin M, Haggerty JL, Lambert M, Poitras ME. Measuring patients' perceptions of patient-centered care: a systematic review of tools for family medicine. Ann Fam Med 2011;9(2):155-164 [FREE Full text] [doi: 10.1370/afm.1226] [Medline: 21403143]
29. Henskens FA, Paul D, Wallis M, Bryant J, Carey M, Fradgley E, et al. Web-based support for population-based medical research: Presenting the QuON Survey System Paper. 2014 Presented at: 7th International Conference on Health Informatics; March 3-6, 2014; Angers, France. [doi: 10.5220/0004738301960204]
30. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174. [Medline: 843571]
31. Flesch R. A new readability yardstick. J Appl Psychol 1948 Jun;32(3):221-233. [Medline: 18867058]
32. Barratt A. How to present the evidence for consumers: preparation of consumer publications. Canberra: Commonwealth of Australia; 2000. URL: http://www.nhmrc.gov.au/_files_nhmrc/publications/attachments/cp66.pdf [accessed 2014-05-15] [WebCite Cache ID 6PZpQDLGZ]
33. Allenby A, Matthews J, Beresford J, McLachlan SA. The application of computer touch-screen technology in screening for psychosocial distress in an ambulatory oncology setting. Eur J Cancer Care (Engl) 2002 Dec;11(4):245-253. [Medline: 12492461]
34. Paul CL, Sanson-Fisher R, Douglas HE, Clinton-McHarg T, Williamson A, Barker D. Cutting the research pie: a value-weighting approach to explore perceptions about psychosocial research priorities for adults with haematological cancers. Eur J Cancer Care (Engl) 2011 May;20(3):345-353 [FREE Full text] [doi: 10.1111/j.1365-2354.2010.01188.x] [Medline: 21414052]

XSL•FO
RenderX

35. Gregory J. Consumer engagement in Australian health policy: final report of the AIHPS research project. Melbourne: Australian Institute of Health Policy Studies; 2008. URL: http://www.healthissuescentre.org.au/documents/items/2009/06/280548-upload-00004.pdf [accessed 2014-05-15] [WebCite Cache ID 6PZpUjl1M]
36. 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]
37. Hepworth J, Robertson AR, Jhunjhunwala A, Jarvis GC, McVittie C. Cancer-related psychosocial research: what are the perspectives of cancer care centre users on participation? Support Care Cancer 2011 Jul;19(7):1029-1035. [doi: 10.1007/s00520-010-0931-9] [Medline: 20563877]
38. Mackenzie LJ, Carey ML, Sanson-Fisher RW, D'Este CA. Psychological distress in cancer patients undergoing radiation therapy treatment. Support Care Cancer 2013 Apr;21(4):1043-1051. [doi: 10.1007/s00520-012-1624-3] [Medline: 23089943]

Edited by G Eysenbach; submitted 14.05.14; peer-reviewed by G Rubin, J Dickerson; comments to author 31.07.14; revised version received 02.09.14; accepted 06.11.14; published 19.12.14.

Please cite as:
Fradgley EA, Paul CL, Bryant J, Roos IA, Henskens FA, Paul DJ
Consumer Participation in Quality Improvements for Chronic Disease Care: Development and Evaluation of an Interactive Patient-Centered Survey to Identify Preferred Service Initiatives
J Med Internet Res 2014;16(12):e292
URL: http://www.jmir.org/2014/12/e292/
doi:10.2196/jmir.3545
PMID:25532217

©Elizabeth A Fradgley, Christine L Paul, Jamie Bryant, Ian A Roos, Frans A Henskens, David J Paul. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 19.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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

Educational Technologies in Problem-Based Learning in Health Sciences Education: A Systematic Review

Jun Jin[1], BA, MA, PhD; Susan M Bridges[2], BA, Dip Ed, GradCert(TESOL), MA (AppLing), EdD

[1]Faculty of Education, The University of Hong Kong, Hong Kong, China

[2]Centre for the Enhancement of Teaching and Learning/Faculty of Education, The University of Hong Kong, Hong Kong, China

Corresponding Author:
Susan M Bridges, BA, Dip Ed, GradCert(TESOL), MA (AppLing), EdD
Centre for the Enhancement of Teaching and Learning/Faculty of Education
The University of Hong Kong
Pokfulam
Hong Kong,
China
Phone: 86 39174771
Fax: 86 25409941
Email: sbridges@hku.hk

Abstract

Background: As a modern pedagogical philosophy, problem-based learning (PBL) is increasingly being recognized as a major research area in student learning and pedagogical innovation in health sciences education. A new area of research interest has been the role of emerging educational technologies in PBL. Although this field is growing, no systematic reviews of studies of the usage and effects of educational technologies in PBL in health sciences education have been conducted to date.

Objective: The aim of this paper is to review new and emerging educational technologies in problem-based curricula, with a specific focus on 3 cognate clinical disciplines: medicine, dentistry, and speech and hearing sciences. Analysis of the studies reviewed focused on the effects of educational technologies in PBL contexts while addressing the particular issue of scaffolding of student learning.

Methods: A comprehensive computerized database search of full-text articles published in English from 1996 to 2014 was carried out using 3 databases: ProQuest, Scopus, and EBSCOhost. Eligibility criteria for selection of studies for review were also determined in light of the population, intervention, comparison, and outcomes (PICO) guidelines. The population was limited to postsecondary education, specifically in dentistry, medicine, and speech and hearing sciences, in which PBL was the key educational pedagogy and curriculum design. Three types of educational technologies were identified as interventions used to support student inquiry: learning software and digital learning objects; interactive whiteboards (IWBs) and plasma screens; and learning management systems (LMSs).

Results: Of 470 studies, 28 were selected for analysis. Most studies examined the effects of learning software and digital learning objects (n=20) with integration of IWB (n=5) and LMS (n=3) for PBL receiving relatively less attention. The educational technologies examined in these studies were seen as potentially fit for problem-based health sciences education. Positive outcomes for student learning included providing rich, authentic problems and/or case contexts for learning; supporting student development of medical expertise through the accessing and structuring of expert knowledge and skills; making disciplinary thinking and strategies explicit; providing a platform to elicit articulation, collaboration, and reflection; and reducing perceived cognitive load. Limitations included cumbersome scenarios, infrastructure requirements, and the need for staff and student support in light of the technological demands of new affordances.

Conclusions: This literature review demonstrates the generally positive effect of educational technologies in PBL. Further research into the various applications of educational technology in PBL curricula is needed to fully realize its potential to enhance problem-based approaches in health sciences education.

(J Med Internet Res 2014;16(12):e251) doi:10.2196/jmir.3240

XSL•FO
RenderX

KEYWORDS

systematic review; educational technologies; problem-based learning; medical education; health sciences; software; digital learning object; interactive whiteboard; learning management system

Introduction

As a modern pedagogical philosophy, problem-based learning (PBL) is increasingly being recognized as a major research area in student learning and pedagogical innovation in health sciences education. In contrast to traditional lecture-dominant teaching and learning approaches, inquiry-based approaches such as PBL prompt students to actively engage in knowledge construction and develop competencies across multiple contexts [1,2]. This review focuses on PBL instead of other distinct inquiry-based pedagogical approaches, such as discovery learning, experiential learning, and project-based learning. Given the high level of technological engagement of 21st century learners, a new area of research interest is examining the role of emerging educational technologies in PBL [3-5]. Therefore, the aim of this paper is to review new and emerging educational technologies in problem-based curricula with a specific focus on 3 cognate disciplines: medicine, dentistry, and speech and hearing sciences. The selection of these 3 related health sciences curricula is based on their level of current activity in the development and research of PBL [6-8]. Of particular interest to this review are studies investigating the role of such technologies in achieving PBL-related student learning outcomes of flexible knowledge, effective problem-solving skills, self-directed learning skills, collaborative teamwork skills, and intrinsic motivation [9,10].

Included studies are ones in which educational technologies have been adopted to support problem-based approaches to learning in both undergraduate and postgraduate programs. The types of technological innovations identified encompass such affordances as learning management system (LMS), specialist learning software (eg, CMapTools), immersive virtual environments (eg, SecondLife), and resources such as 3-dimensional (3D) anatomy models. Also of interest was the use of new hardware, such as interactive whiteboards (IWBs), and how these are combined to reshape new forms of learning in both synchronous, face-to-face "PBL 2.0" [3,11]. Additional studies are exploring the potential to initiate asynchronous models of PBL drawing on distance education needs and modes of delivery [12]. Such innovations draw on the potential of new technologies to provide a rich learning context with access to well-structured information and new spaces for knowledge collaboration [13]. However, although the field is growing and a few reviews have focused on e-learning innovation in health sciences and education [14,15], to date there is no existing systematic review of empirical studies on the usage of educational technologies in PBL in health sciences education.

We have identified 8 roles for technology in learning in the educational literature [16] relevant to identifying studies for inclusion in this review:

1. Access to and structuring of information
2. Curriculum platform
3. Communications media
4. Thinking tools
5. Rich contexts for learning
6. Collaboration spaces
7. A perspective toolkit
8. Scaffolding

The latter issue of scaffolding refers to situations in which experts offer assistance to learners in carrying out new tasks that learners would not be able to complete without support [17]. This issue has been debated in recent PBL and inquiry learning scholarship [18,19] with detractors indicating concerns that PBL does not provide sufficient scaffolding and that the open nature of the problems may add to cognitive load [18]. Proponents argue that PBL is highly scaffolded through strategies such as making disciplinary thinking and strategies explicit, embedding expert guidance, and structuring complex tasks thereby reducing cognitive load [19]. Open to further debate is whether the inclusion of technological affordances such as iPads, laptops, and simulations or variations of synchronous and asynchronous technology-rich delivery of PBL will support or detract from the scaffolding of learning.

Analysis of the studies reviewed will, therefore, focus on the effects of educational technologies in the PBL cycle while addressing the issue of scaffolding of student learning in particular both in face-to-face tutorials and during self-directed learning. The overarching goal is to provide new insights on how learners synthesize information from the multiple technologies employed in PBL at a time of pronounced educational innovation [13,20].

Methods

Focus Questions

Inspired by Cook and West's approach [20] to conducting systematic reviews in medical education and existing review papers [21-24], the screening and classification process conducted is presented subsequently.

The focused questions addressed the population, intervention, comparison, and outcomes (PICO) model as recommended by Cook and West [20]. In addressing the issues above, the research questions addressed in this review are:

1. What effects of educational technologies on students and tutors have been identified in PBL-based applications?
2. How can educational technologies support digitally enhanced and interactive PBL in health sciences education?

Eligibility Criteria

Eligibility criteria for the selection of studies for review were also determined in light of the PICO guidelines. The population was limited to postsecondary education, specifically in dentistry, medicine, and speech and hearing sciences, in which PBL was the key educational pedagogy and curricula. Three types of educational technologies were identified as interventions used to support PBL: learning software and digital learning objects

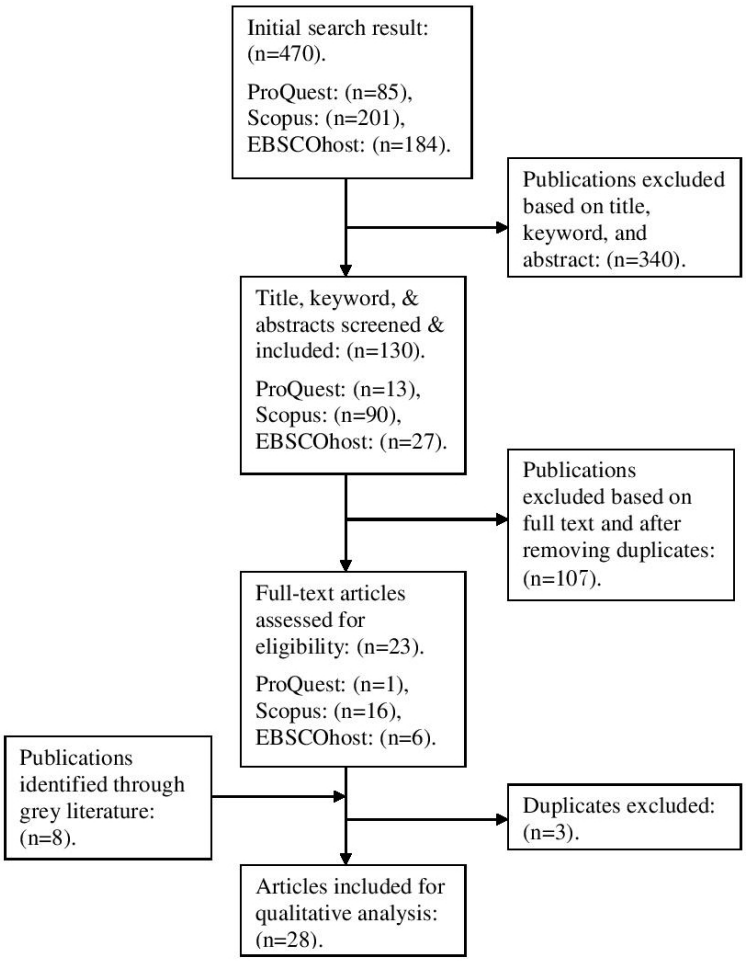
XSL•FO
RenderX

(video/3D models), IWBs and plasma screens, and LMSs. Three types of technologies were selected based on their relatively frequent implementation and innovations as indicated in on-site visitations and communications with health sciences PBL curricula across the globe. Regarding comparisons, although studies adopting experimental designs were included, this was not considered an exclusive criterion given that much educational research in the field is case-based. Finally, included studies indicated outcomes of the effects, both positive and negative, of the use of educational technologies on student learning and staff engagement in PBL. Evidence was determined from both databases and the grey literature.

Selection of Publications

A comprehensive computerized database search of full-text articles published in English from 1996 to 2014 was carried out using 3 education databases: ProQuest, Scopus, and EBSCOhost. Initial search terms were ("educational technologies" OR "learning technologies") AND ("problem-based learning" OR "problem based learning" OR "PBL") AND ("clinical" OR "dent*" OR "med*" OR "speech and hearing"). To narrow down the number of studies retrieved in each database, search terms in title/keywords/abstracts were selected in the initial search stage. The titles and abstracts of retrieved papers were first screened and rated for inclusion based on the PICO inclusion criteria. Additional cross-referencing uncovered grey literature in the form of articles and book chapters. Reviews and commentaries were excluded. The review flowchart (Figure 1) indicates the educational database search method and criteria as well as the final number of studies yielded for analysis (N=28). Search results indicate 3 types of educational technologies, learning software and digital learning objects, IWB (Figure 2 and plasma screens, and LMS, were investigated. Given that LMS combines a range of course or subject management and pedagogical tools to offer a means of designing, building, and delivering online learning environments [25], LMS in the search process includes examples of what are also termed course management systems or CMS (eg, WebCT/Blackboard, Angel, Sakai, and Moodle). Following Cook and West's approach [20], key information (ie, author, year, research design, research purpose, findings) for each article were included. The results were then analyzed and synthesized by narrative or quantitative pooling, exploring effects of educational technologies in PBL-based applications.

Figure 1. Flowchart of the search process.



XSL•FO
RenderX

Figure 2. Photo of an interactive whiteboard.



Results

Overview

Educational technologies have been increasingly used in health sciences education to support or substitute traditional didactic approaches to teaching and learning with inquiry-based approaches. Of the final 28 studies, 20 examined applications of various software and digital learning objects, 5 studies examined the application of IWBs and plasma screens, and 3 explored the application of LMSs in PBL across the 3 clinical disciplines.

Learning Software and Digital Learning Objects

Table 1 indicates studies that have implemented, evaluated, or explored a variety of software and digital learning objects in problem-based clinical education. Software included an interactive distance learning program in obstetrics and gynecology [26], concept mapping [27], 3D visualization [28], CD-ROM (ErgoROM [29]), a text and image database on CD-ROM [30], a Web-based learning portfolio (SkillsBase [31]), online virtual/simulated patients [32-36], video case(s) [37,38], online cases [39,40], an online resource simulating surgical clinical decision making (SURGENT [41]), Interactive Case-based Online Network [42], virtual PBL [43], and computer-based software [44,45]. The majority of these studies analyzed questionnaire data to investigate user perceptions of the efficacy of the various software and digital learning objects that were piloted, implemented, or developed as learning innovations. The purpose of integrating these software and digital learning objects in PBL were reported variously as an aid/supplement [39] or a replacement [34] for traditional formats, such as lectures, dissection, and clinical practice; or for the development of innovative approaches, bridging the gap between theoretical knowledge and clinical practice [29]; or for facilitating collaboration outside of the classroom [42]. Additional implementation goals were the reinforcement of knowledge construction and supporting decision-making processes [33,41], as well as the advancement of teaching and learning [35].

Perceived positive educational impact was seen in providing a more authentic learner environment [33]; conveying and facilitating understanding of information and complex phenomena [28,41]; facilitating enhanced knowledge [35]; improving cognitive, metacognitive, affective, and overall learning processes or outcomes in PBL [37,38]; having a positive impact on active learning [29,34] and critical thinking [29]; a reflective aid to learning clinical skills [31]; providing a suitable environment for collaboration and communication [33,43]; permitting reduced laboratory time; and increasing small-group activity with less reliance on staff [43]. Although the majority were positively disposed toward learning technologies in PBL, 1 study was more critical [32] finding the PBL video scenario to be cumbersome and not imitating real life; therefore, it was seen to be of little educational value.

The key implications include the importance of the modality of the scenario presentation [32] and the need for guiding principles and a direct facilitator connected to the use of 3D visualizations [28]. Hege et al [39] indicated integration of a computer-based learning tool into the curriculum is as important as the optimization of the software itself and concluded that a few aspects or strategies needed to be considered in integration of software into curriculum (eg, the software should be easy-to-use, highly accessible, and should support user evaluation, the delivering of content, user support, and case maintenance). Jha and Duffy [26] proposed 10 “golden rules” from an evaluation of a CD-ROM program in continuing medical education.

XSL•FO
RenderX

Table 1. Studies examining the application of learning software and digital learning objects.

Study	Research design	Research purpose	Findings
Antoniou et al (2014) [36]	The user experienced a specific periodontology case and was asked to provide the optimal responses for each of the challenges of the case. A focus group of 9 undergraduate dentistry students experienced both the Web-based and the Second Life version of this virtual patient.	To assess the suitability of the Second Life MUVE as a virtual patient deployment platform for undergraduate dental education and to explore the requirements and specifications needed to meaningfully repurpose Web-based virtual patients in MUVEs.	The best practices of experiential and immersive game design should be organically incorporated in the repurposing workflow.
Chi et al (2014) [37]	A retrospective cohort study of University of Washington School of Dentistry predoctoral dental students (N=247). The experimental group consisted of first-year dental students (n=63) who received a video case. The historical control group consisted of second-, third-, and fourth-year dental students (n=184) who received a paper case during their first year. A 19-item online survey was administered to all enrolled predoctoral dental students in May 2011.	To compare outcomes associated with video and paper cases used in an introductory public health dentistry course.	Compared to paper cases, video cases significantly improved cognitive, affective, and overall learning outcomes for dental students.
Schwarz et al (2013) [35]	The latest development in the Medical Faculties Network was designed for indexing metadata about simulation-based learning objects. A questionnaire was used to explore students' attitudes and interests in using the interactive algorithms as part of their health sciences studies.	To present and evaluate simulation-based tools for teaching and learning acute medicine issues.	The peer-reviewed algorithms were used for conducting PBL sessions in general medicine and in nursing. The feedback from the survey suggests that the interactive algorithms are effective learning tools to facilitate enhanced knowledge in the field of acute medicine.
Rampling et al (2012) [32]	A total of 24 students tried the scenario and gave feedback via an online survey.	To create a simulated patient with psychosis for psychiatric training within the online virtual environment of Second Life. After design and delivery of the scenario, medical students were asked to complete it and provide feedback.	The scenario was cumbersome, did not imitate real life, and was of little educational value. Multimedia representations of psychotic symptoms were more positively received and there may be scope for further development.
Bridges (2009) [27]	An intervention was conducted with Year 1 students through workshops that introduced the principles of concept mapping software. Qualitative ethnographic data included Year 1 group concept maps and content-specialist feedback on these. Quantitative data were collected using an online questionnaire.	To reports an intervention to improve both the quality of the concept mapping and submission process.	Students have improved in identifying concepts and relationships and that their maps have increased in complexity and are also more clearly presented. After workshops and trialing, concept mapping has positively affected their thinking processes and improved knowledge retention.
Conradi et al (2009) [33]	A virtual patient (VP) scenarios were designed for learners on a Paramedic Foundation Degree within the virtual world Second Life. A player using the MedBiquitous VP international standard allowed cases to be played both within Second Life and on the Web.	To describe and test the PREVIEW project, trailed a replacement to traditional paper PBL with VPs delivered through a virtual world platform.	The Second Life environment engages students effectively in learning, despite some technology barriers. Students perceived SL could provide a more authentic learner environment than classroom-based PBL.
Poulton et al (2009) [34]	A PBL module was converted to VPs, and delivered to 72 students in 10 tutorial groups, with 5 groups each week receiving VPs with options and consequences, and 5 groups receiving online VPs but without options.	To describe and evaluate the replacement of paper cases by interactive online VPs in PBL.	The replacement of paper cases by online interactive VPs was perceived as a success by students and tutors, and both groups appreciated the improvements in PBL process.

XSL•FO
RenderX

Study	Research design	Research purpose	Findings
Corrigan et al (2008) [41]	Evaluation was in both a qualitative and a quantitative fashion. A post-course survey of medical students was used to evaluate access, process, and outcome criteria. SURGENT was used by 98% of students, with 69% spending more than 30 minutes per session on the program.	To assess the introduction of a Web-based module to complement traditional surgical undergraduate curricula.	A Web-enhanced interactive surgical module in an undergraduate course can successfully convey information and understanding beyond the textbook. SURGENT will supplement textbooks and ward experience, allowing students to develop their clinical decision-making skills.
Silen et al (2008) [28]	Questionnaires were used to investigate the medical and physiotherapy students' opinions about the different formats of visualizations and their learning experiences.	To support learning efficacy by developing and using 3D datasets in health sciences curricula and enhancing the knowledge about possible educational value of 3D visualizations in learning anatomy and physiology.	It was successful to implement 3D images in existing themes in programs. The results show that deeper knowledge is required about students' interpretation of images/films regarding learning outcomes. There is also a need for preparations and facilitation principles connected to the use of 3D visualizations.
Hege et al (2007) [39]	Different integration strategies of e-learning are presented and compared regarding motivational aspects and acceptance of students and instructors.	To describe and compare different implementation strategies of case-based learning as an important component of e-learning.	A voluntary integration strategy combined with exam relevance of the content.is recommended. The assets and drawbacks of all described strategies are discussed.
Balslev et al (2005) [38]	11 residents were each assigned to 1 of 2 groups. Both groups analyzed an identical vignette of a patient case. Immediately after, 1 group watched a 2.5-minute video recording and the other group read a description of the same video recording. The groups then reanalyzed the case. Thinking processes were tapped by recording and analyzing the verbal group interaction.	To investigate whether adding a brief video case instead of an equivalent written text improves the cognitive and metacognitive processes of residents in PBL.	The verbal interaction showed statistically significant improvements in data exploration, theory building, and theory evaluation after the video case.
Nathoo et al (2005) [42]	Describe a case study of distinct, small-group tutorials over 2 years as part of the Human Nervous System and Behavior course at the Harvard Medical School. Students and faculty were interviewed following completion of the course and their utilization of the system was recorded and examined.	To assess the introduction of a Web-based innovation in medical education that complements traditional PBL curricula. Utilizing the case method as its fundamental educational approach, the Interactive Case-based Online Network allows students to interact with each other, faculty, and a virtual patient in difficult neurological cases.	This is the first study of the Interactive Case-based Online Network learning system in undergraduate medical education, a platform designed to facilitate collaboration outside of the classroom. Data on user perceptions and system utilization suggest that both faculty and students chose to adopt this online learning system as a means for collaboration.
Hudson (2004) [45]	Third-year medical undergraduates at Adelaide University, South Australia were randomly assigned to 4 groups. Following a pretest, only students in the didactic, problem-based, and free-text groups had 2 weeks of free access to a neuroradiology CAL. Learning was quantified by comparing the post- to pretest scores for each of the 4 groups.	To test the hypothesis that a Computer-aided learning (CAL) tutorial, will result in superior learning (ability to apply and retain knowledge) to that obtained in more passive CAL formats.	While users of an interactive CAL tutorial demonstrated significant learning outcomes compared to non-CAL users, these outcomes were not superior to those achieved from noninteractive CAL.
Sibbald (2004) [40]	Students performed 3 PBL exercises for the same topic. Educational outcomes and students perceptions from an online survey are reported.	To encourage self-directed skill development; addresses learning style preferences; quantitatively and qualitatively assesses the relative advantages of electronic-based vs traditional PBL tools on knowledge and skills building.	This project gave students a balanced, enhanced knowledge perspective from 3 PBL formats; promoted peer teaching, mentoring, and technology skills; and provided insights comparing Web-based tools to other methods for autonomous lifelong learning.
August-Dalfen & Snider (2003) [29]	The program ErgoROM was used with a group of third-year occupational therapy students (n=50) in a semester-long course at McGill University in Canada.	To explore the students' perceptions of using the ErgoROM and impact of the CD-ROM learning experience on their active learning and critical thinking skills.	Overall, 91% of respondents rated the ErgoROM as either "excellent" or "very good." Additionally they reported that ErgoROM had a positive impact on active learning and critical thinking.

XSL•FO
RenderX

Study	Research design	Research purpose	Findings
Bowdish et al (2003) [43]	A quasi experimental, posttest-only research design compared the virtual problem-based learning (VPBL) and a text-based version of the same PBL exercise on students' achievement, as measured by a set of selected physiology examination items, and their perceptions of the learning environment, as measured by the Teaching and Learning Environment Questionnaire.	To report the results and insights of an exploratory investigation of the effectiveness of a prototypic virtual problem-based learning (VPBL) exercise delivered via the Web that uses Hypermedia Assisted Instructional Technologies.	The VPBL is equally as effective as the text-based version for enhancing students' learning and their learning environment in small-group PBL sessions.
Dornan et al (2003) [31]	Qualitative analysis of users' requirements and development of a Web-based learning portfolio. Direct observation of users during a "think-aloud" protocol, a validated software users' measurement inventory, and a questionnaire designed to test whether SkillsBase met its users' requirements.	To evaluate the use of information and communications technology to present a curriculum of clinical skills in a user-friendly format.	SkillsBase meets the design specification for a training and reflective aid to learning clinical skills and is very usable.
Jha & Duffy (2002) [26]	The questionnaire was sent out to 150 Distance Interactive Learning in Obstetrics and Gynaecology (DIALOG). The qualitative data obtained from the evaluation resulted in 10 common items being identified by the majority of respondents.	To carry out a formative evaluation of DIALOG to determine whether DIALOG was achieving its educational objectives.	One of the main benefits of formative assessment is to determine what needs to be done to maintain or improve the program in future. 10 "golden rules" emerged from the evaluation of DIALOG.
Levine et al (1999) [44]	Success of utilization was measured by quantitative improvements in student perceptions and attitudes over a 3-year period.	To describe and assess the implement and integration of computer-based activities into a problem-based gross anatomy curriculum.	The data suggested increasingly positive and beneficial student attitudes toward educational technology, for networks as a faster and more effective method of student/faculty communication, and in the utilization of computer-based instruction for greater flexibility and efficiency in learning the laboratory material.
Andrew & Benbow (1997) [30]	A brief questionnaire was given to a random sample of 100 students at the beginning of the last lecture of the third-year lecture course.	To describe and evaluate the conversion of a traditional image archive into an image resource on compact disc.	This resource useful as an aid to revision, despite relative computer illiteracy, and it is anticipated that students on a new PBL course that incorporates experience with information technology will benefit even more readily when they use the database as an educational resource.

Interactive Whiteboards

Table 2 indicates that the use of IWBs is a new phenomenon in clinical education with 4 studies of IWBs in PBL curricula [3,11,46,47] arising in the past 7 years in addition to 1 study on the use of plasma screens [48] in 2005. Kerfoot [48] indicated that the introduction of computers and plasma screens had a positive impact on PBL tutorials. Bridges and her colleagues [11] adopted an interactional ethnographic methodology to analyze student engagement with digital materials through the use of an IWB and found that "the integration of face-to-face and virtual modalities through the single PBL group's use of an IWB across tutorials and self-study was seamless and supported whole-group engagement in the process." In another study in undergraduate dentistry, Bridges and her colleagues [3] noted that the use of different texts and tools in one problem cycle supported a discursive shift from stimulus for hypothesizing to evidence for final hypotheses. Lu's 2 studies in medical education [46,47] compared a traditional whiteboard group with an IWB group . One study [47] described that nature of scaffolding of collaborative problem solving under the 2 conditions and concluded that educational technology such as IWB can help by expanding the scaffolding choices. The other study [46] identified relationships between learners' collaborative decision making and communicative discourse when engaged in a simulated medical emergency. Group differences were found in that IWB group participants engaged in more adaptive decision-making behavior earlier than the traditional whiteboard group, which led to shared understandings and subsequently to more effective patient management [46]. They also found more productive argumentation in the type of collaborative discourse produced in IWB medical student groups [46]. There are limited studies to show the relative advantages or disadvantages of using IWBs in health sciences education. More studies, therefore, are encouraged to explore both the impact of using IWBs for large-screen visualization and collaboration.

XSL•FO
RenderX

Table 2. Studies examining the application of interactive whiteboards and plasma screens.

Study	Research design	Research purpose	Findings
Bridges et al (2012) [3]	Case study of a single third-year PBL group (n=8) as they engaged in learning activities across a problem cycle. The study investigated the data trail across PBL learning events and contexts and the various discourse members. One discourse member of the PBL group was selected as an anchor point for tracking across the data collection.	To investigate PBL-as-process in clinical education through detailed analysis of the "way" students and their tutors construct knowledge and negotiate meaning in situ in a dental PBL curriculum.	The way students experience and understand 2 "black box" facets of their PBL learning, independent learning, and online learning, are explored by adopting interactional ethnography multimodality within a theory of semiotics examines multimodal texts support cognition and transformative learning.
Bridges et al (2010) [11]	IWBs were installed in all PBL tutorial rooms and IT support was provided for all students and facilitators in a 5-year undergraduate dental PBL curriculum. Year 1 students (n=55) received additional workshops and IWB support. Analysis of 4 hours of video-recorded learning activities undertaken by a Year 1 undergraduate PBL group (n=8) on 3 occasions during the same week.	To establish and evaluate how the progression from a digital repository approach toward an interactive blending of technology within face-to-face tutorials might be supported and resourced.	The group accessed a range of digital materials to support learning within and across all phases of a problem cycle. The use of learning objects and online resources within a problem cycle supported inquiry learning and the discursive shift in student talk from a stimulus for hypothesizing in relatively lay terms in the first tutorial to evidence. The integration of face to-face and virtual modalities through the use of an IWB within the tutorial was seamless and supported whole-group engagement in the problem process.
Lu et al (2010) [47]	2 conditions (traditional whiteboard and IWB) using a "deteriorating patient" case were examined. 2 groups of third-year medical students in the Department of Internal Medicine in a large urban teaching hospital volunteered to participate.	To describe the nature of scaffolding of collaborative problem solving under 2 conditions—with technological support and without.	Although appropriate scaffolding is still based on the teacher's domain knowledge and pedagogy experience, technology can help by expanding the scaffolding choices that an instructor can make in a medical training context.
Lu & Lajoie (2008) [46]	2 groups of 7 third-year medical students volunteered to participate. Both groups were asked to solve a learning activity. 1 group of students used a traditional whiteboard and the other used an IWB while solving the patient problem. 2 kinds of data were collected: collaborative decision-making discourse and computer records of whiteboard annotations.	To investigate the collaborative decision making and communicative discourse of groups of learners engaged in a simulated medical emergency across the 2 condition subgroups.	IWB enabled data sharing and construction of shared understandings about the patient. Shared visualization clarified verbal interaction, promoted productive argumentation, and facilitated negotiation. Argumentation tools embedded into the IWB design enhanced groups' decision making and communicative interactions in the simulated medical emergency.
Kerfoot et al (2005) [48]	37 tutorial groups, were observed to record the patterns of use of the computers and plasma screens. Based on these observations, surveys were developed and distributed to students and tutors.	To examine how the introduction of this educational technology impacted PBL tutorials.	Both students and tutors reported that the introduction of computers and wall-mounted plasma screens had positively impacted tutorials. Questions were raised as to how this technology might alter tutorial dynamics.

Learning Management Systems

Table 3 lists studies that adopted various LMSs to support problem-based approaches. These included the application of WebCT [49], VMS Portal [50] to manage inquiry-based materials and activities for PBL curricula, and iSUS for self-directed learning [51].

A few studies noted the positive effects of using LMS in PBL curricula [49-51]. In Dornan's study [51], a Web-based LMS-ISUS helped to provide practical guidance about what to learn and how to learn, helped access appropriate experiences and manage time, gave feedback on students' accumulated real patient learning, provided peer comparison, and helped self-management. The authors argued that ISUS can provide the motivational jigsaw to fill the gap between PBL and placement learning. McLean [49] found students perceived that LMS are the most useful means of communication and resource delivery for PBL in medicine. His study also highlighted limitations in that insufficient support, resources, and training might result in less successful implementation of educational technologies.

With regard to the question of scaffolding, some of the previous studies noted that monitoring, support, and development are important for efficient and positive implementation of an LMS in PBL curricula. In the VMS Portal project [50], medical students were involved in website development to help, consolidate, integrate, and develop Web resources for their peers. Critical to successful scaffolding in PBL are tutor facilitation strategies [10]. In online environments, tutor presence, ongoing engagement, and timely feedback become factors in facilitating students' problem solving, self-directed learning, and collaboration in health sciences education.

XSL•FO
RenderX

Table 3. Studies examining the application of learning management systems.

Study	Research design	Research purpose	Findings
Rosenbaum et al (2009) [50]	VMS Portal used software and database technology for a highly customized Web portal for medical students. Access to course material, evaluations, academic information, and community assets were customized for individual users. Modular features were added in response to student requests and feedback and monitoring of usage habits.	To describe a practice-based focus by 2 medical students to create a website for all medical students.	Medical students are uniquely positioned to help consolidate, integrate, and develop Web resources for peers. As other medical schools create and expand digital resources, input by medical students should be solicited.
Dornan et al (2005) [51]	A Web-based LMS included 66 placement students in a PBL medical curriculum. Data were free-response comments from 16 students during 7 weeks of usage, transcripts of pre- and postgroup discussions, and questionnaire responses (100%).	To establish whether and under what conditions medical students can learn in a self-directed manner in the clinical environment.	Students valued affective and pedagogic support, and relied on teachers to manage their learning environment. With support, they were motivated and able to choose how and when to meet their learning needs.
McLean & Murrell (2002) [49]	WebCT in PBL, student-centered curriculum was introduced. A survey was conducted after completing the first module in School of Medicine in South Africa.	To gather user feedback with regard to the value of WebCT as a curriculum support, especially the value of WebCT for the delivery of digitized material.	Students responded positively to the communication facility. WebCT will be particularly useful when students are off campus, undertaking electives and community service.

Discussion

Principal Findings

The journal articles and book chapters examined in this systematic review indicate the generally positive effect of the thoughtful implementation of educational technologies in PBL. This is particularly the case where such technologies support scaffolding thereby reducing cognitive load and allowing students to learn in complex domains [19]. Firstly, when considering resource development, educational technologies enable provision of rich, authentic problems and/or case contexts accessible on demand in virtual spaces. Online virtual/simulated patients, video case(s), and online cases [33-41] convey complex phenomena in a more authentic learning environment. Secondly, educational technologies provide not only access to engagement in the problem-based inquiry, but also structure information by embedding expert knowledge and skills. This may be in the form of problem-relevant videos and simulations made available during self-directed learning [3] demonstrating case reports independent of time and place. Thirdly, educational technologies support students and their facilitators in making disciplinary thinking explicit. Dedicated software can help learners to construct explanations, structure tasks, and make them more manageable [19]. Adaptations of standard LMS as well as dedicated inquiry-based LMS can provide a platform to elicit articulation, collaboration, and reflection [52]. Technological resources, such as different software and 3D models, LMS, and IWB, can assist students engage in problem-solving processes.

Although generally positive, a limited number of studies have indicated the adverse effects of educational technologies or their methods of implementation. In this review paper, the less successful implementations may be attributable to the content or delivery of the video scenario [32]. Additional cognitive burden due to higher levels of complexity was seen as a possible limitation to their effective implementation; however, in highly positive cases, technologies were seen as providing an additional supportive scaffold for student learning. In terms of infrastructure and staff development, insufficient support, resources, and training [49,53] were seen as disabling. Adopting and/or adapting functionalities from software or LMS with limited staff-student and/or student-student interactivity and limited student feedback processes were seen as shortcomings in cases where this occurred. This would indicate that thoughtful instructional design approaches need to be applied when adapting more traditional systems or when designing new programs for inquiry-based learning [54]. Well-designed empirical research studies are needed to establish best practices for technological hardware and software in enhancing teaching and learning productivity and building stronger learning communities [55].

Although the major forms of educational technologies presented in these studies look very promising and are potentially fit for educational purpose in many problem-based health sciences education settings, there is a need for support and training in light of the ever-changing nature of both technical knowledge of teaching staff and the technological affordances themselves. The power of computers and Internet, merging with multimedia and interactive spaces, appear to allow a high degree of flexibility and accessibility to digital instructional content in health sciences education [56]. These affordances may enable integration of technologies into the curriculum and support staff development by encouraging interactive teaching and learning. However, the often-false assumption is the “sink or swim” approach in which faculty may assume students’ prior technological skills or knowledge. Where new affordances are introduced or even old affordances with new inquiry-based purposes, additional hands-on tutorials and training are helpful to facilitate student adaptation to the technology and ensure optimal benefit [49,57]. Successful multimedia teaching and

XSL•FO
RenderX

learning relies not only on the proper use of information technology, but also on a clear implementation strategy [58]. Implications also arise in considering Technological Pedagogical Content Knowledge [59] with regard to disciplinary influences on the ways that educational technologies are incorporated into the curriculum. Guidance and support should be tailored to meet the needs of each user including learner, facilitators, and curriculum developers to be at its most effective.

Conclusions

In conclusion, this literature review indicates a generally positive effect from the adoption of various educational technologies in PBL. Positive outcomes for student learning included providing rich, authentic problems and/or case contexts for learning; supporting student development of medical expertise through the accessing and structuring of expert knowledge and skills; making disciplinary thinking and strategies explicit; providing a platform to elicit articulation, collaboration, and reflection; and reducing perceived cognitive load. Insufficient technical support, infrastructure, and resources were seen as impacting negatively on uptake and learning outcomes. Staff and student induction and ongoing training in the use of educational technologies for learning in inquiry-based contexts such as PBL is recommended.

Although educational technologies have been increasingly used in health sciences education, it has been questioned whether they can completely substitute traditional teaching methods [60]. The rise of Massive Online Open Courses in all fields, including health sciences, has been seen as positive, particularly for continuous medical education and public health literacy [15]. In considering undergraduate inquiry-based curricula, this review supports Hmelo-Silver [19] and Bridges et al's [61] predictions that technology can play an important but synergistic role with other components of PBL. Further research into the various applications of educational technologies in PBL curricula is needed to fully realize their potential in enhancing inquiry-based approaches in health sciences education. In an increasingly digital, networked world, convergence of educational technologies is increasingly apparent. This has given rise to understandings that learners are positioned within digital ecosystems. Consequently, it is possible that a learner might engage with the merging of distinct educational technologies. The effects of learning in a digital ecosystem need to be identified and explored in further research.

Acknowledgments

This work was partially funded by The University of Hong Kong Sciences of Learning Strategic Research Theme. Research assistance provided by Ms Rita PC Suen, Ms Jenny Lee, and Dr Alexandra Chong Boon Hor was greatly appreciated. The authors particularly acknowledge the invaluable contribution of Professor Tara Whitehill to the foundation of this project and dedicate this paper to her memory.

Conflicts of Interest

None declared.

References

1. Prosser M, Sze D. Problem-based learning: student learning experiences and outcomes. Clin Linguist Phon 2014 Jan;28(1-2):131-142. [doi: 10.3109/02699206.2013.820351] [Medline: 23944271]
2. Walker A, Leary H. A problem based learning meta analysis: differences across problem types, implementation types, disciplines, and assessment levels. Interdisciplinary Journal of Problem-based Learning 2009 Mar 24;3(1):12-43. [doi: 10.7771/1541-5015.1061]
3. Bridges SM, Botelho MG, Green J, Chau ACM. Multimodality in problem-based learning (PBL): an interactional ethnography. In: Bridges SM, McGrath CP, Whitehill T, editors. Problem-Based Learning in Clinical Education: the Next Generation. Netherlands: Springer; 2012:99-120.
4. Bridges S, Green J, Botelho MG, Tsang PCS. Blended learning and PBL: An interactional ethnographic approach to understanding knowledge construction in-situ. In: Walker A, Leary H, Hmelo-Silver C, Ertmer PA, editors. Essential Readings in Problem-Based Learning. West Lafayette, IN: Purdue University Press; 2015.
5. Lu J, Bridges S, Hmelo-Silver CE. Problem-based learning. In: Sawyer K, editor. Cambridge Handbook of the Learning Sciences. Cambridge: Cambridge University Press; 2014.
6. Schmidt HG, Rotgans JI, Yew EH. The process of problem-based learning: What works and why. Medical Education 2011;45(8):792-806. [doi: 10.1111/j.1365-2923.2011.04035]
7. Shuler CF. Comparisons in basic science learning outcomes between students in PBL and traditional dental curricula at the same dental school. In: Bridges S, McGrath C, Whitehill TL, editors. Problem-Based Learning in Clinical Education: The Next Generation. Netherlands: Springer; 2012.
8. Whitehill TL, Bridges S, Chan K. Problem-based learning (PBL) and speech-language pathology: a tutorial. Clin Linguist Phon 2014 Jan;28(1-2):5-23. [doi: 10.3109/02699206.2013.821524] [Medline: 23944231]
9. Hmelo-Silver CE. Problem-based learning: what and how do students learn? Educational Psychology Review 2004 Sep;16(3):235-266. [doi: 10.1023/B:EDPR.0000034022.16470.f3]

XSL•FO
RenderX

10. Hmelo-Silver CE, Barrows HS. Goals and strategies of a problem-based learning facilitator. Interdisciplinary Journal of Problem-based Learning 2006 May 22;1(1):21-39. [doi: 10.7771/1541-5015.1004]
11. Bridges SM, Botelho MG, Tsang PCS. PBL.2.0: Blended learning for an interactive, problem-based pedagogy. Medical Education 2010;44(11):1131. [doi: 10.1111/j.1365-2923.2010.03830.x]
12. Ng ML, Bridges S, Law SP, Whitehill T. Designing, implementing and evaluating an online problem-based learning (PBL) environment--a pilot study. Clin Linguist Phon 2014 Jan;28(1-2):117-130. [doi: 10.3109/02699206.2013.807879] [Medline: 23837407]
13. Hmelo-Silver CE, Eberbach C. Learning theories and problem-based learning. In: Bridges SM, McGrath CP, Whitehill T, editors. Problem-Based Learning in Clinical Education: The Next Generation. Netherlands: Springer; 2012:3-17.
14. Hansen MM. Versatile, immersive, creative and dynamic virtual 3-D healthcare learning environments: a review of the literature. J Med Internet Res 2008;10(3):e26 [FREE Full text] [doi: 10.2196/jmir.1051] [Medline: 18762473]
15. Liyanagunawardena TR, Williams SA. Massive open online courses on health and medicine: Review. Journal of Medical Internet Research 2014;16(8):e191. [doi: 10.2196/jmir.1051]
16. Goldman‐Segall R, Maxwell JW. Computers, the Internet, and new media for learning. In: Reynolds WM, Miller GE, editors. Handbook of Psychology. Volume 7: Educational Psychology. New York: John Wiley & Sons; 2002:393-427.
17. Rogoff B, Gardner WP. Adult guidance of cognitive development. In: Everyday Cognition: Its Development in Social Context. Cambridge, MA: Harvard University Press; 1984:95-116.
18. Kirschner PA, Sweller J, Clark RE. Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. Educational Psychologist 2006 Jun;41(2):75-86. [doi: 10.1207/s15326985ep4102_1]
19. Hmelo-Silver CE, Duncan RG, Chinn CA. Scaffolding and achievement in problem-based and inquiry learning: a response to Kirschner, Sweller, and Clark (2006). Educational Psychologist 2007 Apr 26;42(2):99-107. [doi: 10.1080/00461520701263368]
20. Cook DA, West CP. Conducting systematic reviews in medical education: a stepwise approach. Med Educ 2012 Oct;46(10):943-952. [doi: 10.1111/j.1365-2923.2012.04328.x] [Medline: 22989128]
21. Needleman IG. A guide to systematic reviews. Journal of Clinical Periodontology 2002;29(s3):6-9. [doi: 10.1034/j.1600-051X.29.s3.15.x]
22. Steinert Y, Mann KV. Faculty development: principles and practices. Journal of Veterinary Medical Education 2011 Feb 10;33(3):317-324. [doi: 10.3138/jvme.33.3.317]
23. Kuijpers W, Groen WG, Aaronson NK, van Harten WH. A systematic review of web-based interventions for patient empowerment and physical activity in chronic diseases: relevance for cancer survivors. J Med Internet Res 2013 Feb;15(2):e37 [FREE Full text] [doi: 10.2196/jmir.2281] [Medline: 23425685]
24. 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;15(4):e85 [FREE Full text] [doi: 10.2196/jmir.1933] [Medline: 23615206]
25. Sleator RD. The evolution of eLearning background, blends and blackboard. Sci Prog 2010;93(Pt 3):319-334. [Medline: 21047021]
26. Jha V, Duffy S. 'Ten golden rules' for designing software in medical education: results from a formative evaluation of DIALOG. Med Teach 2002 Jul;24(4):417-421. [doi: 10.1080/01421590220145798] [Medline: 12193327]
27. Bridges SM, Dyson JE, Corbet EF. Blended learning, knowledge co-construction and undergraduate group work. Medical Education 2009;43(5):490-491. [doi: 10.1111/j.1365-2923.2009.03345.x]
28. Silén C, Wirell S, Kvist J, Nylander E, Smedby O. Advanced 3D visualization in student-centred medical education. Med Teach 2008 Jun;30(5):e115-e124. [doi: 10.1080/01421590801932228] [Medline: 18576181]
29. August-Dalfen S, Snider L. A multimedia case based approach to the study of office ergonomics. Work 2003;20(1):3-11. [Medline: 12631997]
30. Andrew SM, Benbow EW. Conversion of a traditional image archive into an image resource on compact disc. Journal of Clinical Pathology 1997 Jul 01;50(7):544-547. [doi: 10.1136/jcp.50.7.544]
31. Dornan T, Maredia N, Hosie L, Lee C, Stopford A. A web-based presentation of an undergraduate clinical skills curriculum. Med Educ 2003 Jun;37(6):500-508. [doi: 10.1046/j.1365-2923.2003.01531.x]
32. Rampling J, O'Brien A, Hindhaugh K, Woodham L, Kavia S. Use of an online virtual environment in psychiatric problem-based learning. The Psychiatrist 2012 Oct 01;36(10):391-396. [doi: 10.1192/pb.bp.111.037630]
33. Conradi E, Kavia S, Burden D, Rice A, Woodham L, Beaumont C, et al. Virtual patients in a virtual world: Training paramedic students for practice. Med Teach 2009 Aug;31(8):713-720. [Medline: 19811207]
34. Poulton T, Conradi E, Kavia S, Round J, Hilton S. The replacement of 'paper' cases by interactive online virtual patients in problem-based learning. Medical Teacher 2009;31(8):752-758. [doi: 10.1007/s10459-009-9154-6]
35. Schwarz D, Štourač P, Komenda M, Harazim H, Kosinová M, Gregor J, et al. Interactive algorithms for teaching and learning acute medicine in the network of medical faculties MEFANET. J Med Internet Res 2013 Jul;15(7):e135 [FREE Full text] [doi: 10.2196/jmir.2590] [Medline: 23835586]

XSL•FO
RenderX

36. Antoniou PE, Athanasopoulou CA, Dafli E, Bamidis PD. Exploring design requirements for repurposing dental virtual patients from the web to second life: a focus group study. J Med Internet Res 2014;16(6):e151 [FREE Full text] [doi: 10.2196/jmir.3343] [Medline: 24927470]
37. Chi DL, Pickrell JE, Riedy CA. Student learning outcomes associated with video vs. paper cases in a public health dentistry course. J Dent Educ 2014 Jan;78(1):24-30. [Medline: 24385521]
38. Balslev T, de Grave WS, Muijtjens AM, Scherpbier AJ. Comparison of text and video cases in a postgraduate problem-based learning format. Med Educ 2005 Nov;39(11):1086-1092. [doi: 10.1111/j.1365-2929.2005.02314.x] [Medline: 16262803]
39. Hege I, Ropp V, Adler M, Radon K, Mäsch G, Lyon H, et al. Experiences with different integration strategies of case-based e-learning. Med Teach 2007 Oct;29(8):791-797. [doi: 10.1080/01421590701589193] [Medline: 18236274]
40. Sibbald D. A student assessment of the virutal interactive case tool for asynchronous learning (VITAL) and other self-directed learning formats. American Journal of Pharmaceutical Education 2004;68(1):1-7.
41. Corrigan M, Reardon M, Shields C, Redmond H. "SURGENT" -- student e-learning for reality: the application of interactive visual images to problem-based learning in undergraduate surgery. J Surg Educ 2008 Mar;65(2):120-125. [doi: 10.1016/j.jsurg.2007.11.011] [Medline: 18439533]
42. Nathoo AN, Goldhoff P, Quattrochi JJ. Evaluation of an Interactive Case-based Online Network (ICON) in a problem based learning environment. Adv Health Sci Educ Theory Pract 2005 Aug;10(3):215-230. [doi: 10.1007/s10459-005-7851-3] [Medline: 16193402]
43. Bowdish B, Chauvin SW, Kreisman N, Britt M. Travels towards problem based learning in medical education (VPBL). Instructional Science 2003;31(4):231-253. [doi: 10.1023/A:1024625707592]
44. Levine MG, Stempak J, Conyers G, Walters JA. Implementing and integrating computer-based activities into a problem-based gross anatomy curriculum. Clinical Anatomy 1999;12(3):191-198. [doi: 10.1002/(SICI)1098-2353(1999)12:3<191::AID-CA8>3.0.CO;2-5]
45. Hudson JN. Computer-aided learning in the real world of medical education: does the quality of interaction with the computer affect student learning? Med Educ 2004 Aug;38(8):887-895. [doi: 10.1111/j.1365-2929.2004.01892.x] [Medline: 15271050]
46. Lu J, Lajoie SP. Supporting medical decision making with argumentation tools. Contemporary Educational Psychology 2008;33(3):425-442. [doi: 10.1016/j.cedpsych.2008.05.005]
47. Lu J, Lajoie SP, Wiseman J. Scaffolding problem-based learning with CSCL tools. Computer Supported Learning 2010 Jul 10;5(3):283-298. [doi: 10.1007/s11412-010-9092-6]
48. Kerfoot BP, Masser BA, Hafler JP. Influence of new educational technology on problem-based learning at Harvard Medical School. Med Educ 2005 Apr;39(4):380-387. [doi: 10.1111/j.1365-2929.2005.02105.x] [Medline: 15813760]
49. McLean M, Murrell K. WebCT: integrating computer-mediated communication and resource delivery into a new problem-based curriculum. J Audiov Media Med 2002 Mar;25(1):8-15. [doi: 10.1080/01405110220118365] [Medline: 12025491]
50. Rosenbaum BP, Gorrindo TL, Patel SG, McTigue MP, Rodgers SM, Miller BM. Medical student involvement in website development. Med Teach 2009 Jul;31(7):627-633. [Medline: 19811147]
51. Dornan T, Hadfield J, Brown M, Boshuizen H, Scherpbier A. How can medical students learn in a self-directed way in the clinical environment? Design-based research. Med Educ 2005 Apr;39(4):356-364. [doi: 10.1111/j.1365-2929.2005.02112.x] [Medline: 15813757]
52. Quintana C, Reiser BJ, Davis EA, Krajcik J, Fretz E, Duncan RG, et al. A scaffolding design framework for software to support science inquiry. Journal of the Learning Sciences 2004 Jul;13(3):337-386. [doi: 10.1207/s15327809jls1303_4]
53. Reynolds PA, Eaton KA, Mason R. Supporting the learner and teacher online. Br Dent J 2008 Jun 14;204(11):625-629. [doi: 10.1038/sj.bdj.2008.462] [Medline: 18552803]
54. Dørup J. Experience and attitudes towards information technology among first-year medical students in Denmark: longitudinal questionnaire survey. J Med Internet Res 2004 Mar 5;6(1):e10 [FREE Full text] [doi: 10.2196/jmir.6.1.e10] [Medline: 15111276]
55. Kamel Boulos M, Wheeler S. The emerging Web 2.0 social software: an enabling suite of sociable technologies in health and health care education. Health Info Libr J 2007 Mar;24(1):2-23. [doi: 10.1111/j.1471-1842.2007.00701.x] [Medline: 17331140]
56. Mattheos N, Schoonheim-Klein M, Walmsley AD, Chapple IL. Innovative educational methods and technologies applicable to continuing professional development in periodontology. Eur J Dent Educ 2010 May;14 Suppl 1:43-52. [doi: 10.1111/j.1600-0579.2010.00624.x] [Medline: 20415976]
57. Crowe P, Howie C, Thorpe M. Self-reported computer literacy of medical students and staff: good facilities are not enough. Med Teach 1998 Jan;20(3):264-265. [doi: 10.1080/01421599881057]
58. Yip HK, Barnes IE. Information technology: information technology in dental education. Br Dent J 1999 Sep 25;187(6):327-332. [doi: 10.1038/sj.bdj.4800271]
59. Mishra P, Koehler MJ. Technological pedagogical content knowledge: a framework for teacher knowledge. Teachers College Record 2006;108(6):1017-1054.

XSL•FO
RenderX

60. Venail F, Deveze A, Lallemant B, Guevara N, Mondain M. Enhancement of temporal bone anatomy learning with computer 3D rendered imaging software. Med Teach 2010 Jul;32(7):e282-e288. [doi: 10.3109/0142159X.2010.490280] [Medline: 20653370]
61. Bridges SM, McGrath C, Whitehill T. The next generation: research directions in PBL. In: Bridges SM, McGrath C, Whitehill T, editors. Problem-Based Learning in Clinical Education: The Next Generation. Netherlands: Springer; 2012:225-231.

Abbreviations

3D: 3-dimensional
CAL: computer-aided learning
CMS: course management systems
DIALOG: Distance Interactive Learning in Obstetrics and Gynaecology
IWB: interactive whiteboard
LMS: learning management system
PBL: problem-based learning
PICO: population, intervention, comparison, and outcomes
VP: virtual patient
VPBL: virtual problem-based learning

Edited by G Eysenbach; submitted 12.01.14; peer-reviewed by G Finger, T Klobučar; comments to author 26.04.14; revised version received 16.06.14; accepted 17.08.14; published 10.12.14.

Please cite as:
Jin J, Bridges SM
Educational Technologies in Problem-Based Learning in Health Sciences Education: A Systematic Review
J Med Internet Res 2014;16(12):e251
URL: http://www.jmir.org/2014/12/e251/
doi:10.2196/jmir.3240
PMID:25498126

©Jun Jin, Susan M Bridges. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 10.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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

Personal Experiences and Emotionality in Health-Related Knowledge Exchange in Internet Forums: A Randomized Controlled Field Experiment Comparing Responses to Facts Vs Personal Experiences

Joachim Kimmerle[1], Dipl-Psych, Dr rer nat; Martina Bientzle[1], Dipl-Psych; Ulrike Cress[1], Dipl-Psych, Dr rer soc

Knowledge Media Research Center (KMRC), Knowledge Construction Lab, Tübingen, Germany

Corresponding Author:
Joachim Kimmerle, Dipl-Psych, Dr rer nat
Knowledge Media Research Center (KMRC)
Knowledge Construction Lab
Schleichstr 6
Tübingen, 72076
Germany
Phone: 49 7071 979 363
Fax: 49 7071 979 100
Email: j.kimmerle@iwm-kmrc.de

Abstract

Background: On the Internet, people share personal experiences as well as facts and objective information. This also holds true for the exchange of health-related information in a variety of Internet forums. In online discussions about health topics, both fact-oriented and strongly personal contributions occur on a regular basis.

Objective: In this field experiment, we examined in what way the particular type of contribution (ie, factual information vs personal experiences) has an impact on the subsequent communication in health-related Internet forums.

Methods: For this purpose, we posted parallelized queries to 28 comparable Internet forums; queries were identical with regard to the information contained but included either fact-oriented descriptions or personal experiences related to measles vaccination. In the factual information condition, we posted queries to the forums that contained the neutral summary of a scientific article. In the personal experiences condition, we posted queries to the forums that contained the same information as in the first condition, but were framed as personal experiences

Results: We found no evidence that personal experiences evoked more responses (mean 3.79, SD 3.91) from other members of the Internet forums than fact-oriented contributions (mean 2.14, SD 2.93, t_{26}=0.126, P=.219). But personal experiences elicited emotional replies (mean 3.17, SD 1.29) from other users to a greater extent than fact-oriented contributions (mean 2.13, SD 1.29, t_{81}=3.659, P<.001).

Conclusions: We suggest that personal experiences elicited more emotional replies due to the process of emotional anchoring of people's own style of communication. We recommend future studies should aim at testing the hypotheses with more general and with less emotionally charged topics, constructing different fact-oriented posts, and examining additional potential factors of influence such as personality factors or particular communication situations.

(J Med Internet Res 2014;16(12):e277) doi:10.2196/jmir.3766

KEYWORDS

personal experiences; emotionality; health; knowledge exchange; field experiment; Internet forums

Introduction

The Internet is an essential platform for sharing and exchanging information on a multitude of topics [1,2]. There is, however, great variety in the types of information that are shared on the Internet, ranging from purely fact-oriented information exchange to sharing very personal information [3-5]. People with a common interest in a particular topic may come together on the Internet as online communities in order to share information [6,7]. Platforms that are particularly relevant for information

XSL•FO
RenderX

exchange are Internet forums [8,9]. These forums are virtual places where users can ask questions, express their opinions, and share their knowledge and experiences [10]. Participants in Internet forums can post messages that other users can pick up and reply to—instantly or at a later point in time. Many forums are organized and arranged according to certain topics or subtopics [8,10].

One subject area for which the Internet is an important platform for information search and information exchange is the field of health-related issues [11,12]. In the exchange of health-related information in pertinent Internet forums [5,10,13,14] and online health communities [15-18], there is also a wide range in the variety of contributions, with both fact-oriented information exchange and the transfer of personal information. Health-related Internet forums make it possible for everyone to make contributions, that is, ask or answer questions.

At the same time, this type of very open information exchange involves the risk that the reliability of health information is often very unclear [19-22]. Moreover, users who ask questions in an Internet forum are often not aware of the kinds of replies they can expect. They also cannot know under which circumstances their chances of receiving an answer are good or bad, if any answer at all [23]. Should they make their request as matter-of-factly, objectively, and neutrally as possible? Or do queries have a higher chance of eliciting an answer when they express personal concerns and experiences? And beyond the pure frequency of answers, what will reply posts look like then in terms of personalization and emotionalization? Will objective or personal queries elicit, in turn, answers that are respectively more neutral or more emotional?

Previous research has shown that people are very interested in personal information and care a lot about communication that includes personal and emotional aspects [4,24-26]. In particular, personal experience is considered a highly legitimate form of giving advice regarding health-related issues [26-28]. Accordingly, we state the following hypothesis: Hypothesis 1: users will reply more frequently to health-related queries that consist of personal experiences than to corresponding queries that contain fact-oriented information.

In addition, research has found that humans tend to imitate other people's behavior [29] and the communication style of their communication partners [30-32]. In particular, people have a propensity to imitate expressions of emotion [33-35]. Moreover, emotionally expressive people were found to be more cooperative than people with a low level of emotional expressivity [36]. Accordingly, we state an additional hypothesis: Hypothesis 2: users will reply in a more emotional manner to health-related queries that consist of personal experiences than to corresponding queries that contain fact-oriented information.

Methods

Sample

We selected 28 Internet forums for this study. We took into consideration six different types of forums with regard to topic: 8 forums had a focus on health, 8 on family issues, 2 were sports forums, 2 dealt with naturopathy, 2 were vaccination critique forums, and 6 were general forums that dealt with a wide variety of topics including sports, health, etc.

From those considered, we selected only forums in which posts had occurred on the day of selection. Moreover, we ensured that the forums selected all had a comparable level of activity. A complete list of all 28 forums is included in Multimedia Appendix 1.

Procedure

We applied a procedure that allowed for a randomized and controlled approach in combination with a high level of ecological validity. We created two experimental conditions: (1) factual information, and (2) personal experiences. In the factual information condition, we posted queries to the forums that contained the neutral summary of a scientific article. In the personal experiences condition, we posted queries to the forums that contained the same information as in the first condition, but were framed as personal experiences (see below). We randomly assigned an equal number of the different types of forums to the two experimental conditions, resulting in an allocation of 4 health forums, 4 family forums, 1 sports forum, 1 naturopathy forum, 1 vaccination critique forum, and 3 general forums in each condition.

We created user accounts for all forums on the same day and with the same gender-neutral user name. These accounts were used to post the queries in the respective forums on the identical day. We recorded the answers to these posts in the subsequent 4 weeks. In order to maintain absolute standardization of the experimental procedure, we did not reply to any answers or counter questions.

Experimental Material

Both queries dealt with the topic of measles vaccination. We chose this topic because vaccination is a heavily debated issue about which many people have strong opinions and tend to express them on the Internet [37,38]. Accordingly, it appeared to be plausible that a substantial number of forum users would be willing to reply to such queries and express their opinion on this topic.

The query in the factual information condition was developed on the basis of an article published in the official journal of the German Medical Association [39]. This article reported the outbreak of measles in a medium-sized town and provided statistical descriptions of this epidemic. The query that was posted in this condition indicated that this user was looking for opinions about measles vaccination and had already gathered some information. This introduction was followed by a brief summary of the journal article including the complete reference. The provision of the reference was intended to emphasize the pure fact-orientation of this query and to make clear to the recipients that this information did not result from any kind of personal experience. This post consisted of 87 words.

The query in the personal experiences condition also stated that the user was looking for opinions about measles vaccination and had previously gathered some information. This introduction was followed by the same information as in the previous

XSL•FO
RenderX

condition (regarding the age of the children concerned and the symptoms), but was framed in terms of personal experience with the outbreak of measles in a day-care center. This post did not refer to the journal article and consisted of 69 words.

Measures

We counted the number of replies and captured the average number of replies to the posts in the two experimental conditions in order to test Hypothesis 1. In addition, two independent raters evaluated the emotionality of the replies on 5-point Likert scales ranging from 1 (neutral) to 5 (very emotional), in order to allow for examining Hypothesis 2. Those raters were blind to the respective experimental condition and to the general research questions, and reached consensus on their final rating in cases of disagreement.

The following reply is an example of a very emotional statement:

> *For me, measles vaccination is obligatory and I was really glad that my little kid could finally be vaccinated now in May! Because I am extremely frightened of the long-term consequences and I personally rather put up with the very unlikely vaccination adverse effects than running the risk of a measles infection. Once in the hospital I had to care for a child suffering from meningitis, which died then, and this experience has probably affected me …*

This reply, in contrast, is an example of an emotionally neutral statement:

> *Yes, measles are contagious, yes, children can fall ill with them, and yes, one can be vaccinated against them, which is worthwhile, because measles are known for having an impact on fertility.*

Statistical Analysis

We compared the number of replies within the different types of forums using chi-square tests. In order to test Hypothesis 1, we compared the number of replies with a chi-square test and conducted an independent samples *t* test with the average number of replies as dependent variable. To take the distribution of this variable into account, we additionally conducted a Mann-Whitney *U* test and a Mood's median test. For testing Hypothesis 2, we conducted an independent samples *t* test with emotionality of the replies as dependent variable.

Ethical Considerations

Since this study used information posted in open online forums, contributions were considered as belonging to the public domain [40,41]. The posts of all users in all forums could be openly read by everybody; usernames and passwords were only required for posting. So the behavior of research participants did not occur in a private context [42]. In line with the requirements of the local ethics committee at the Knowledge Media Research Center (Tübingen, Germany), the dignity and integrity of participants were not violated in any way in this study. We did not give medical advice to the forum users. We did not provide any medically questionable or incorrect information in the online forums.

We made sure that our queries did not confront the users with any kind of information that they would not otherwise encounter in Internet forums or many other communication situations. Thus, the forum users were not taken advantage of in any respect; there was absolutely no infringement of the users' rights. In order to protect the individual user's personal or online identity, we do not provide information about the names or online identities of contributors in this article. In order to avoid an individual's identity being traced by searching for the quoted phrases, we use only English translations of direct quotes [42,43].

Results

Number of Replies

There was a wide range in the number of replies among the different Internet forums: the number of replies per forum varied from 0 to 13. The most replies were found in the 8 family forums (a total of n=37; mean 4.63, SD 3.89) compared to the other 20 forums (a total of n=46; mean 2.30, SD 3.18), as indicated by a chi-square test (χ^2_{10}=22.12, *P*=.015). The other forum types did not differ among each other in the frequency of replies, all *P*s>.370.

Regarding Hypothesis 1, there appeared to be more replies (n=53) to the personal experiences than to the factual information (n=30), with a median of 2 in the personal experiences condition and a median of 1 in the factual information condition. But the two conditions did not differ significantly from each other when we compared the number of replies with a chi-square test (χ^2_{10}=8.40, *P*=.590).

The same finding occurred when we conducted an independent samples *t* test with the average number of replies as dependent variable (t_{26}=1.26, *P*=.219). In the factual information condition, the posts received on average 2.14 (SD 2.93) replies per forum. In the personal experiences condition, we found 3.79 (SD 3.91) replies.

We also conducted further nonparametric tests, but both a Mann-Whitney *U* test (*P*=.178) and a Mood's median test (*P*=.257) confirmed that the difference was non-significant.

Emotionality of Replies

As expected in Hypothesis 2, in a *t* test the replies of the other forum users were significantly more emotional in the personal experiences condition (mean 3.17, SD 1.29) than the answers in the factual information condition (mean 2.13, SD 1.29; t_{81}=3.66, *P*<.001).

Discussion

Principal Findings

Contrary to our initial assumptions, personal experiences did not elicit more replies than factual information. This absence of a significant effect was probably due to the very large variance of replies in the different forums. It might also be the case that forum users tended to regard factual information and personal experience as equivalent. This interpretation would be

XSL•FO
RenderX

consistent with previous findings that forum users in health-related contexts do not adequately distinguish between verifiable and non-verifiable information [10,16,21,22].

The key finding of this field experiment is that answers to a post containing personal experiences were significantly more emotional than replies to factual information. Reasons for this could be that inquirers who reported private information were perceived as more personable or that the recipients of the initial posts felt addressed on a personal level and, accordingly, replied on the same level. This behavior might be explained by effects of anchoring and adjusting [44-46], in that participants used the initial post as an emotional anchor to which they adjusted their own style of communication.

Limitations and Future Work

The finding that reply rates were rather low overall could be due to the topic we selected. Potentially, and contrary to our expectations that people would feel strongly about the topic, the subject area of measles vaccination is only important to a limited group of people, that is, people who have young children themselves. It may be a rather abstract or even an irrelevant topic for childless people, such as singles or adolescents. This interpretation is supported by the finding that the family forums were the most active ones regarding our queries, compared to forums concerned with other topics. In addition, most answers implied that they were written by people who had personal experience with this topic. This might also indicate that vaccination is in fact an emotionally charged topic for many people, a circumstance that may have had an impact on their response behavior. Future studies should aim to test the hypotheses with different, that is, more general topics with a broader base of interest and with a less emotionally charged subject matter.

Another limitation of our study is that the queries in the factual information condition might have been perceived to be quite uncommon or peculiar, because it is rather seldom that posts in Internet forums include a reference to a journal article. Even though the provision of this reference was intended to underline the fact-orientation of the post, it is not clear how the forum users perceived this information. Future research should try to find ways to construct fact-oriented posts that do not appear to be peculiar to other forum users.

In addition, the emotionality of the reply posts was only evaluated by two independent raters. Future studies might consider applying additional assessment methods, such as identifying key words of emotionality in the reply post of the forum users [47,48].

Finally, it might be worthwhile for further research to examine additional potential factors of influence that may have an impact on the way people express their answers. This may include certain personality factors, such as social value orientation [49], agreeableness [50], or people's need for self-presentation and self-esteem [51,52]. In addition, these factors of influence may also include aspects of the communication situation, such as anonymity [53], identifiability [54-56], or the synchronicity [57-59] of information exchange. The current study represents only a beginning to potential research into what impact particular kinds of contributions have on the subsequent exchanges of health-related information in Internet forums.

Acknowledgments

We would like to thank the participants of the Knowledge Communication course within the Psychology degree program at the University of Tübingen in the summer term 2012 for contributing to the preparation of the stimulus material and the collection of the data.

Conflicts of Interest

None declared.

Multimedia Appendix 1

List of Internet forums.

[PDF File (Adobe PDF File), 14KB - jmir_v16i12e277_app1.pdf]

References

1. Bargh JA, McKenna KY. The internet and social life. Annu Rev Psychol 2004;55:573-590. [doi: 10.1146/annurev.psych.55.090902.141922] [Medline: 14744227]
2. Joinson A, McKenna KYA, Postmes T, Reips UD. The Oxford handbook of Internet psychology. Oxford: Oxford University Press; 2007.
3. Barnett S, Jones SC, Caton T, Iverson D, Bennett S, Robinson L. Implementing a virtual community of practice for family physician training: a mixed-methods case study. J Med Internet Res 2014;16(3):e83 [FREE Full text] [doi: 10.2196/jmir.3083] [Medline: 24622292]
4. Humphreys L, Gill P, Krishnamurthy B. Twitter: a content analysis of personal information. Information, Communication & Society 2013 Oct 18;17(7):843-857. [doi: 10.1080/1369118X.2013.848917]
5. Stjernswärd S, Hansson L. A web-based supportive intervention for families living with depression: content analysis and formative evaluation. JMIR Res Protoc 2014;3(1):e8 [FREE Full text] [doi: 10.2196/resprot.3051] [Medline: 24550185]

XSL•FO
RenderX

6. Preece J. Sociability and usability in online communities: Determining and measuring success. Behaviour & Information Technology 2001 Jan;20(5):347-356. [doi: 10.1080/01449290110084683]
7. Preece J, Maloney-Krichmar D. Online communities: Design, theory and practice. J Comput Mediat Commun 2005;10(4):1. [doi: 10.1111/j.1083-6101.2005.tb00264.x]
8. Holtz P, Kronberger N, Wagner W. Analyzing internet forums. Journal of Media Psychology: Theories, Methods, and Applications 2012 Aug 31;24(2):55-66. [doi: 10.1027/1864-1105/a000062]
9. Sudau F, Friede T, Grabowski J, Koschack J, Makedonski P, Himmel W. Sources of information and behavioral patterns in online health forums: observational study. J Med Internet Res 2014;16(1):e10 [FREE Full text] [doi: 10.2196/jmir.2875] [Medline: 24425598]
10. Kimmerle J, Gerbing KK, Cress U, Thiel A. Exchange of complementary and alternative medical knowledge in sport-related Internet fora. Sociol Sport J 2012;29(3):348-364.
11. Lima-Pereira P, Bermúdez-Tamayo C, Jasienska G. Use of the Internet as a source of health information amongst participants of antenatal classes. J Clin Nurs 2012 Feb;21(3-4):322-330. [doi: 10.1111/j.1365-2702.2011.03910.x] [Medline: 22093043]
12. McInnes N, Haglund BJ. Readability of online health information: implications for health literacy. Inform Health Soc Care 2011 Dec;36(4):173-189. [doi: 10.3109/17538157.2010.542529] [Medline: 21332302]
13. Hua M, Alfi M, Talbot P. Health-related effects reported by electronic cigarette users in online forums. J Med Internet Res 2013;15(4):e59 [FREE Full text] [doi: 10.2196/jmir.2324] [Medline: 23567935]
14. Keeling D, Khan A, Newholm T. Internet forums and negotiation of healthcare knowledge cultures. Journal of Services Marketing 2013 Feb 18;27(1):59-75. [doi: 10.1108/08876041311296383]
15. Aarts JW, Faber MJ, den Boogert AG, Cohlen BJ, van der Linden PJ, Kremer JA, et al. Barriers and facilitators for the implementation of an online clinical health community in addition to usual fertility care: a cross-sectional study. J Med Internet Res 2013 Aug;15(8):e163 [FREE Full text] [doi: 10.2196/jmir.2098] [Medline: 23996964]
16. Kimmerle J, Thiel A, Gerbing K, Bientzle M, Halatchliyski I, Cress U. Knowledge construction in an outsider community: Extending the communities of practice concept. Computers in Human Behavior 2013 May;29(3):1078-1090. [doi: 10.1016/j.chb.2012.09.010]
17. Zhao J, Ha S, Widdows R. Building trusting relationships in online health communities. Cyberpsychol Behav Soc Netw 2013 Sep;16(9):650-657. [doi: 10.1089/cyber.2012.0348] [Medline: 23786170]
18. Zaslow E. Revalorizing feminine ways of knowing: the challenge to biomedical epistemology in an online mothers' health community. Inf Commun Soc 2012 Dec;15(9):1352-1372. [doi: 10.1080/1369118X.2011.638653]
19. Craigie M, Loader B, Burrows R, Muncer S. Reliability of health information on the Internet: an examination of experts' ratings. J Med Internet Res 2002;4(1):e2 [FREE Full text] [doi: 10.2196/jmir.4.1.e2] [Medline: 11956034]
20. Hardey M. Doctor in the house: the Internet as a source of lay health knowledge and the challenge to expertise. Sociol Health & Illness 1999 Nov;21(6):820-835. [doi: 10.1111/1467-9566.00185]
21. Miah A, Rich E. The medicalization of cyberspace. Oxon & New York: Routledge; 2008.
22. Rains SA. The anonymity effect: the influence of anonymity on perceptions of sources and information on health websites. Journal of Applied Communication Research 2007 May;35(2):197-214. [doi: 10.1080/00909880701262666]
23. Gabbiadini A, Mari S, Volpato C. Virtual users support forum: do community members really want to help you? Cyberpsychol Behav Soc Netw 2013 Apr;16(4):285-292. [doi: 10.1089/cyber.2012.0412] [Medline: 23530547]
24. Colombo C, Mosconi P, Confalonieri P, Baroni I, Traversa S, Hill SJ, et al. Web search behavior and information needs of people with multiple sclerosis: focus group study and analysis of online postings. Interact J Med Res 2014 Jul;3(3):e12 [FREE Full text] [doi: 10.2196/ijmr.3034] [Medline: 25093374]
25. Greene DD, Heeteer C. Personal stories within virtual environments: a cancer patient information software case study. CyberPsychology & Behavior 1998 Jan;1(3):201-211. [doi: 10.1089/cpb.1998.1.201]
26. Maloney P. Online networks and emotional energy: how pro-anorexic websites use interaction ritual chains to (re)form identity. Information, Communication & Society 2013 Feb;16(1):105-124. [doi: 10.1080/1369118X.2012.659197]
27. McCosker A, Darcy R. Living with cancer: Affective labour, self-expression and the utility of blogs. Information, Communication & Society 2013 Oct;16(8):1266-1285. [doi: 10.1080/1369118X.2012.758303]
28. Sillence E. Giving and receiving peer advice in an online breast cancer support group. Cyberpsychol Behav Soc Netw 2013 Jun;16(6):480-485. [doi: 10.1089/cyber.2013.1512] [Medline: 23659724]
29. Chartrand TL, Lakin JL. The antecedents and consequences of human behavioral mimicry. Annu Rev Psychol 2013 Jan;64:285-308. [doi: 10.1146/annurev-psych-113011-143754] [Medline: 23020640]
30. Moore R. Imitation and conventional communication. Biol Philos 2012 Oct 20;28(3):481-500. [doi: 10.1007/s10539-012-9349-8]
31. Seitz S, Stewart C. Imitation and expansions: some developmental aspects of mother-child communications. Developmental Psychology 1975;11(6):763-768. [doi: 10.1037/0012-1649.11.6.763]
32. Shea N. Imitation as an inheritance system. Philos Trans R Soc Lond B Biol Sci 2009 Aug 27;364(1528):2429-2443 [FREE Full text] [doi: 10.1098/rstb.2009.0061] [Medline: 19620113]
33. Häfner M, Ijzerman H. The face of love: spontaneous accommodation as social emotion regulation. Pers Soc Psychol Bull 2011 Dec;37(12):1551-1563. [doi: 10.1177/0146167211415629] [Medline: 21778468]

XSL•FO
RenderX

34. van der Velde SW, Stapel DA, Gordijn EH. Imitation of emotion: When meaning leads to aversion. Eur J Soc Psychol 2009;40(3):536-542. [doi: 10.1002/ejsp.667]
35. Wallbott HG. Recognition of emotion from facial expression via imitation? Some indirect evidence for an old theory. Br J Soc Psychol 1991 Sep;30(3):207-219. [doi: 10.1111/j.2044-8309.1991.tb00939.x]
36. Schug J, Matsumoto D, Horita Y, Yamagishi T, Bonnet K. Emotional expressivity as a signal of cooperation. Evolution and Human Behavior 2010 Mar;31(2):87-94. [doi: 10.1016/j.evolhumbehav.2009.09.006]
37. Betsch C, Sachse K. Debunking vaccination myths: strong risk negations can increase perceived vaccination risks. Health Psychol 2013 Feb;32(2):146-155. [doi: 10.1037/a0027387] [Medline: 22409264]
38. Zimmerman RK, Wolfe RM, Fox DE, Fox JR, Nowalk MP, Troy JA, et al. Vaccine criticism on the World Wide Web. J Med Internet Res 2005 Jun;7(2):e17 [FREE Full text] [doi: 10.2196/jmir.7.2.e17] [Medline: 15998608]
39. Arenz S, Kalies H, Ludwig MS, Hautmann W, Siedler A, Liebl B, et al. Der Masernausbruch in Coburg [The measles epidemic in Coburg]. Deutsches Ärzteblatt 2003 Jul;100(49):3245-3249.
40. Hewitt-Taylor J, Bond CS. What e-patients want from the doctor-patient relationship: content analysis of posts on discussion boards. J Med Internet Res 2012 Dec;14(6):e155 [FREE Full text] [doi: 10.2196/jmir.2068] [Medline: 23137788]
41. Hookway N. 'Entering the blogosphere': some strategies for using blogs in social research. Qualitative Research 2008 Feb 01;8(1):91-113. [doi: 10.1177/1468794107085298]
42. Eysenbach G, Till JE. Ethical issues in qualitative research on internet communities. BMJ 2001 Nov 10;323(7321):1103-1105. [doi: 10.1136/bmj.323.7321.1103]
43. Snee H. Web 2.0 as a social science research tool. 2008. URL: http://www.restore.ac.uk/orm/futures/Web_2.0_final__v3.pdf [accessed 2014-11-23] [WebCite Cache ID 6UISkfNk3]
44. Kimmerle J, Cress U. The impact of cognitive anchors on information-sharing behavior. Cyberpsychol Behav Soc Netw 2013 Jan;16(1):45-49. [doi: 10.1089/cyber.2012.0144] [Medline: 23153079]
45. Mussweiler T, Strack F. The semantics of anchoring. Organizational Behavior and Human Decision Processes 2001 Nov;86(2):234-255. [doi: 10.1006/obhd.2001.2954]
46. Mussweiler T, Strack F. Hypothesis-consistent testing and semantic priming in the anchoring paradigm: a selective accessibility model. Journal of Experimental Social Psychology 1999 Mar;35(2):136-164. [doi: 10.1006/jesp.1998.1364]
47. Strapparava C, Mihalcea R. Learning to identify emotions in text. In: Proceedings of the 2008 ACM Symposium on Applied Computing. 2008 Presented at: ACM Symposium on Applied Computing; March 16-20, 2008; Fortaleza, Ceará, Brazil.
48. Strapparava C, Mihalcea R. Annotating and identifying emotions in text. In: Intelligent Information Access. Berlin, Heidelberg: Springer; 2010:21-38.
49. Kimmerle J, Wodzicki K, Jarodzka H, Cress U. Value of information, behavioral guidelines, and social value orientation in an information-exchange dilemma. Group Dynamics: Theory, Research, and Practice 2011;15(2):173-186. [doi: 10.1037/a0021467]
50. Landers RN, Lounsbury JW. An investigation of Big Five and narrow personality traits in relation to Internet usage. Computers in Human Behavior 2006 Mar;22(2):283-293. [doi: 10.1016/j.chb.2004.06.001]
51. Rui JR, Stefanone MA. Strategic image management online: Self-presentation, self-esteem and social network perspectives. Information, Communication & Society 2013 Oct;16(8):1286-1305. [doi: 10.1080/1369118X.2013.763834]
52. van Uden-Kraan CF, Drossaert CH, Taal E, Seydel ER, van de Laar MA. Self-reported differences in empowerment between lurkers and posters in online patient support groups. J Med Internet Res 2008 Jun;10(2):e18 [FREE Full text] [doi: 10.2196/jmir.992] [Medline: 18653442]
53. Frost J, Vermeulen IE, Beekers N. Anonymity versus privacy: selective information sharing in online cancer communities. J Med Internet Res 2014;16(5):e126 [FREE Full text] [doi: 10.2196/jmir.2684] [Medline: 24828114]
54. Cress U, Kimmerle J. Endowment heterogeneity and identifiability in the information-exchange dilemma. Computers in Human Behavior 2008 May;24(3):862-874. [doi: 10.1016/j.chb.2007.02.022]
55. Douglas KM, McGarty C. Internet identifiability and beyond: A model of the effects of identifiability on communicative behavior. Group Dynamics: Theory, Research, and Practice 2002;6(1):17-26. [doi: 10.1037/1089-2699.6.1.17]
56. Douglas KM, McGarty C. Identifiability and self-presentation: computer-mediated communication and intergroup interaction. Brit J Soc Psychol 2001 Sep;40(3):399-416. [doi: 10.1348/014466601164894]
57. Cress U, Kimmerle J, Hesse FW. Impact of temporal extension, synchronicity, and group size on computer-supported information exchange. Computers in Human Behavior 2009 May;25(3):731-737. [doi: 10.1016/j.chb.2009.01.010]
58. Helquist JH, Krus J, Nunamaker Jr JF. Synchronicity and group ability to autonomously cluster brainstorming ideas. Int J e-Collab 2009;5(4):67-81. [doi: 10.4018/jec.2009062605]
59. Swaab RI, Galinsky AD, Medvec V, Diermeier DA. The communication orientation model: explaining the diverse effects of sight, sound, and synchronicity on negotiation and group decision-making outcomes. Pers Soc Psychol Rev 2012 Feb;16(1):25-53. [doi: 10.1177/1088868311417186] [Medline: 21846835]

XSL•FO
RenderX

Edited by G Eysenbach; submitted 07.08.14; peer-reviewed by P Talbot, H Potts, N Walton; comments to author 11.09.14; revised version received 27.09.14; accepted 06.11.14; published 04.12.14.

Please cite as:
Kimmerle J, Bientzle M, Cress U
Personal Experiences and Emotionality in Health-Related Knowledge Exchange in Internet Forums: A Randomized Controlled Field Experiment Comparing Responses to Facts Vs Personal Experiences
J Med Internet Res 2014;16(12):e277
URL: http://www.jmir.org/2014/12/e277/
doi:10.2196/jmir.3766
PMID:25486677

©Joachim Kimmerle, Martina Bientzle, Ulrike Cress. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 04.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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

Which Frail Older Patients Use Online Health Communities and Why? A Mixed Methods Process Evaluation of Use of the Health and Welfare Portal

Peter Makai[1], MSc, PhD; Marieke Perry[1], MD, PhD; Sarah HM Robben[2], MD, PhD; Henk Schers[1], MD, PhD; Maud Heinen[1], RN, PhD; Marcel GM Olde Rikkert[1], MD, PhD; René JF Melis[1], MD, PhD

[1]Radboud University Medical Center, Radboud Institute for Health Sciences, Nijmegen, Netherlands

[2]Jeroen Bosch Hospital, Department of Internal Medicine, 's-Hertogenbosch, Netherlands

Corresponding Author:
Peter Makai, MSc, PhD
Radboud University Medical Center
Radboud Institute for Health Sciences
Renier Postlaan 4
Nijmegen, 6500 HB
Netherlands
Phone: 31 24 36 19807
Fax: 31 024 361 74 08
Email: Peter.Makai@radboudumc.nl

Abstract

Background: Frail older people often receive fragmented care from multiple providers. According to the literature, there is an urgent need for coordination of care. Online and eHealth tools are increasingly used to improve coordination. However, there are significant barriers to their implementation in frail older people.

Objective: Our aim was to (1) evaluate differences in use of a personal online health community (POHC) for frail older people in relation to personal characteristics, and (2) explore barriers and facilitators for use as experienced by older people and their informal caregivers, using the case of the Health and Welfare Information Portal (ZWIP).

Methods: This is a mixed methods study. For the quantitative analysis, we used POHC usage information (2 years follow-up) and baseline characteristics of frail older people. For the qualitative analysis, we used semistructured interviews with older people and their informal caregivers. Participants were recruited from 11 family practices in the east of the Netherlands and frail older people over 70 years. The ZWIP intervention is a personal online health community for frail older people, their informal caregivers, and their providers. ZWIP was developed at the Geriatrics Department of Radboud University Medical Center. We collected data on POHC use for 2 years as well as relevant patient characteristics. Interview topics were description of use, reasons for use and non-use, and user profiles.

Results: Of 622 frail patients in the intervention group, 290 were connected to ZWIP; 79 used ZWIP regularly (at least monthly). Main predictors for use were having an informal caregiver, having problems with activities of daily living, and having a large number of providers. Family practice level predictors were being located in a village, and whether the family practitioners had previously used electronic consultation and cared for a large percentage of frail older people. From 23 interviews, main reasons for use were perceiving ZWIP to be a good, quick, and easy way of communicating with providers and the presence of active health problems. Important reasons for non-use were lack of computer skills and preferring traditional means of consultation.

Conclusions: Only 27.2% (79/290) of frail older enrolled in the POHC intervention used the POHC frequently. For implementation of personal online health communities, older people with active health problems and a sizable number of health care providers should be targeted, and the informal caregiver, if present, should be involved in the implementation process.

Trial Registration: International Standard Randomized Controlled Trial Number (ISRCTN): 11165483; http://www.controlled-trials.com/isrctn/pf/11165483 (Archived by WebCite at http://www.webcitation.org/6U3fZovoU).

(J Med Internet Res 2014;16(12):e278) doi:10.2196/jmir.3609

XSL•FO
RenderX

KEYWORDS

eHealth; frail older people; care-coordination; chronic care

Introduction

Potential Benefits of eHealth for Frail Older People

Frail older people have a large number of health deficits and receive fragmented care, often resulting in nursing home admissions and hospitalization [1]. Frail older people also have multiple health care providers. The collaboration between these providers is not always optimal, and they are often unaware of the fact that they care for the same patient. This reveals insufficient coordination in the care for frail older people [2]. There are several barriers to providing coordinated care to frail older people in a cross-institutional setting: (1) the physical barriers of working in different institutions located in different geographical locations, (2) not having a timely overview of all professionals involved in the care of the patient, (3) information that colleagues have on the patient, (4) how they treat the patient, and (5) that multidisciplinary consultation can take place only if all professionals are available at the same time. The required coordination of care for frail older people may be facilitated by the implementation of eHealth [3,4].

Personal online health communities (POHCs) seem particularly suited to improving the coordination of care for frail older people with multiple providers. Such communities allow a patient, and/or professionals caring for a specific patient, to interact in online networks. In general, online health communities are online platforms that unite providers and/or patients with a shared goal or similar interest [5]; in the case of POHCs, that is the care for an individual. POHCs enable communication between people who would not have met each other otherwise [5]. They also provide an alternative to face-to-face consultations. POHCs facilitate communication by organizing a network of providers around a patient [5]. They are a relatively new development in eHealth [6], and so information on the determinants of usage is limited. We developed the Health and Welfare Information Portal (in Dutch: Zorg- en WelzijnsInfoPortaal or ZWIP), a POHC aimed at reducing fragmentation in care in frail older people. ZWIP targets individual patients, their informal caregivers, and their health care providers, to learn about the feasibility and effects of POHCs in the care for frail older people. In an earlier study, we showed that ZWIP had not yet yielded any benefits [7], but there were indications that usage was heterogeneous across users and, in general, limited. Usage is therefore the focus of this study.

Determinants of Using Personal Online Health Communities in Older People

Despite the potential benefits of POHCs, implementing eHealth is difficult, especially in frail elderly populations [3]. Implementation difficulties in this group arise because frail older people have decreased physical and cognitive function leading to a loss of autonomy. Older people in general have a lower level of computer literacy than younger populations, which has to be considered when developing and implementing eHealth interventions for this group [8]. Furthermore, older patients may not be able to use complex and multifunctional eHealth interventions, which also leads to low levels of usage.

At present, little is known about facilitators of eHealth usage in general and POHCs in particular in frail older people and how the barriers presented can be overcome [3]. Previous studies focused on implementation barriers during the introduction of eHealth interventions in other groups, such as health care providers [9-14] and in younger patients with chronic illnesses [15-17]. Knowledge about such determinants of usage in frail older people would allow identification of older people who are likely to use and possibly benefit from eHealth interventions.

Aim

The aim of this study was to investigate the main determinants of ZWIP usage in frail older people. First, we describe ZWIP use and its users' characteristics. Second, we investigate the predictors (barriers and facilitators) of ZWIP usage in frail older people to identify successful users.

Methods

Intervention

ZWIP is a personal online health community for multidisciplinary communication and information exchange for frail older people and their informal caregivers. The development of the intervention has been described elsewhere [2]. Briefly, ZWIP is a secure digital environment, aiming to improve collaboration within the care network around frail older people. Within the ZWIP network, frail older people (or on their request, their informal caregiver) take the lead. Patients and their informal caregivers give permission to providers to join their network. For all members this creates an overview of the providers involved in their care. ZWIP contains a message system similar to email, where patients can exchange messages with these providers. Most messages are visible to all members of a patient's network, with the exception of private messages. Therefore, all relevant professionals and informal caregivers are kept up to date on care-related developments. Based on the family practice's information system and frailty identification through Easycare-TOS (see below), ZWIP contains current medical and social care data on frail older people, which is shared within a patient's network. ZWIP also gives the opportunity to register individual patient's care-related goals and action plans. Goals and action plans are determined by the patients in consultation with the professionals. Patients may also receive tailored health information via ZWIP. (See Multimedia Appendix 1 for a video of a patient using ZWIP.)

Centers, Participants, and Recruitment

ZWIP was implemented in 11 family practices in the east of the Netherlands. The participating practices started by screening all of their patients aged 70 years and older in alphabetical order using the Easycare-TOS (Two step Older persons Screening) instrument. Easycare-TOS is a validated, two-step measure to identify frail older people in the community as a target population for integrated geriatric care [18]. Easycare-TOS is

XSL•FO
RenderX

based on the EASYcare-system, a focused geriatric assessment by a primary care nurse [19,20]. Initially, the family practitioner completed a questionnaire in Step 1 and judged patients as frail, frailty uncertain, or not frail based on readily available information. If patients were assessed as being frail or family practitioners were uncertain, a second assessment step (Step 2) was carried out by specialized nurses to confirm or exclude frailty. All 622 patients who were considered frail based on the screening were invited to participate in ZWIP, of which 290 (46.6%) consented.

Design

In complex interventions, which are usually difficult to describe and replicate, qualitative research can help clarify the processes of implementation and change. Therefore, we evaluated ZWIP usage with a mixed methods study. The local ethics committee (Committee on Research Involving Human Subjects Region Arnhem-Nijmegen) stated that no formal approval was required due to the non-invasive and non-experimental nature of the study [21,22].

Quantitative Data Collection

Quantitative data were collected in the intervention arm of an effectiveness trial [7], and measurements were performed by trained nurses in the patients' homes using a face-to-face questionnaire. We used baseline data for demographics, physical, mental and social functioning, as well as coordination of care and care use. We included patient characteristics observed by the family practitioner or obtained during the Easycare-TOS assessment. Individual baseline characteristics used were the Barthel index of Activities of Daily Living (ADL) and the Lawton scale of Instrumental Activities of Daily Living (IADL), the Short Form-36 Health Survey (SF-36) (social activity limitations and mental health) [23-26], as well as education, age, and sex. Family practice level characteristics were practice location, family practitioner having used electronic consultation previously, and percentage of frail older people in the practice. Furthermore, we used data from the ZWIP application's logging files. In this model, ZWIP usage was operationalized as monthly total page views by the patients or their informal caregivers. ZWIP use was followed up to 2 years. We operationalized ZWIP usage as a combined measure of total page views by the patients or their informal caregivers. To capture all forms of usage, we counted passive usage (simply logging in without interaction with the providers, viewing health promotion) and active usage by the older people or their informal caregivers (actual communication with providers, adapting goals).

Quantitative Data Analysis

Baseline characteristics of ZWIP participants and non-participants were compared using *t* tests and chi-square tests. We classified different groups of users according to usage of ZWIP, using hierarchical cluster analysis [27], creating groups from observations based on similarities of their features. ZWIP users were clustered according to the degree of use of the various ZWIP features (communication, goals, health promotion, etc) [27]. Clusters were created using the Ward method [28], and the number of separate groups were determined using the pseudo t^2 and the pseudo F^2 [27,29].

Furthermore, we investigated the relationship between usage and baseline characteristics at the patient and family practice levels. We used a hierarchical linear growth model [30] to account for potential variation in use across time (months since connection to ZWIP) within a patient, between patients with various characteristics, and between family practices. All analyses were performed using SAS 9.2.

Qualitative Data Collection

Participants were patients and their informal caregivers who participated in ZWIP and were selected using purposive sampling [31]. To include a wide variety of users and levels of use during sampling, we allowed for the following patient characteristics: sex, age, family practice, whether a patient or an informal caregiver managed the patient's ZWIP account, network size, and how much they used ZWIP (ZWIP usage). If any patient declined to participate, a patient with comparable characteristics was chosen instead. The final sample size was dictated by data saturation [32].

Three research assistants, who were not otherwise involved in the evaluation of ZWIP, approached the selected patients or their informal caregivers and conducted individual semistructured interviews at the patients' homes. The interviewers used a topic list based on relevant themes derived from the literature [33], which evolved when relevant new topics came up. Initial topics were description of use, reasons for use/non-use, and user profiles. New topics were discussed in the research group and pursued if relevant to the aim of this study. All interviews were audiotaped and transcribed verbatim. Data collection continued until saturation was reached (eg, when no new themes emerged from the interviews).

Qualitative Data Analysis

To analyze the transcripts, 3 researchers (PM, SR, and MP) used directed content analysis [34] and deconstructed the interviews independently, using the principles of iterative comparison. In discussion meetings, PM, SR, and MP reached agreement on codes, combined codes to generate and adjust categories (axial coding), merged categories to compose and refine themes (selective coding), and discussed whether saturation had been reached. We used Atlas-ti 5.2 software to help in data coding and retrieval.

Integration of Quantitative and Qualitative Data

Quantitative data were used to identify qualitative respondents. We based the user profiles on the clusters identified during cluster analysis, thus making the link between qualitative and quantitative data. Qualitative and quantitative data on facilitators for use were presented separately and integrated by triangulation [35] in the discussion. Similar weight was given to the two types of analyses.

XSL•FO
RenderX

Results

Participants

A total of 290 vulnerable older patients (46.6% of all 622 eligible patients) completed baseline measurement and participated in ZWIP. The other 332 refused participation. ZWIP participants were similar to non-users in terms of demographic criteria, functional measures such as ADL and mental health measures, number of illnesses, and the types of illnesses present with the exception of dementia. Furthermore, they were similar in terms of a number of process indicators such as patient experience with care measures, and family practitioner reported process measures, with one exception: patients using ZWIP were more likely to have more providers involved in their care (Table 1).

In total, 23 patients and informal caregivers participated in the interviews. Patient age varied between 74 and 90 years, and they came from a wide range of practices (see Multimedia Appendix 2). Additionally, as a result of purposive sampling, interviewed patients varied substantially in terms of messages sent and received and if they had an informal caregiver and the type of informal caregivers (spouse or child).

Usage

The most frequently used ZWIP function was messaging with an average of 12 messages per patient during the 2-year follow-up period (Table 2). Additionally, 47.9% (139/290) of patients used the goal-setting function and 13.1% (38/290) of older people modified or evaluated the goals during ZWIP usage. Furthermore, 33.1% (96/290) of patients had defined concrete care-related activities to reach their goals; in the case of 9.0% (26/290) of patients, these activities were modified or actively evaluated in ZWIP during the period of ZWIP usage.

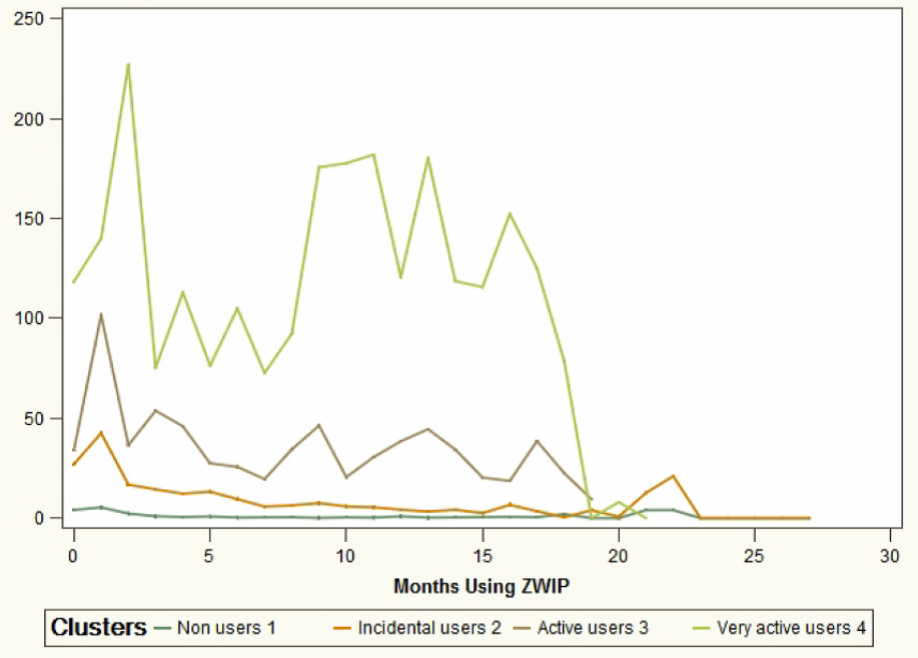
Hierarchical cluster analysis identified four main groups of users: Non-active users (using ZWIP only a few times within 2 years), Regular users (once a month), Active users (once a week), and a small group of Very active users (daily users). From quantitative demographic data and qualitative data, average user profiles per cluster were constructed (Textbox 1). As can be seen in Figure 1, after an enthusiastic start the number of page views decreased to remain relatively stable in all groups a few months after the start, except for the very active users (Figure 1).

In addition, qualitative data further characterized the variations in use seen in the quantitative data trends. Qualitative data confirmed that patients and informal caregivers made highly variable use of ZWIP: “I check ZWIP every day” [Respondent 21, Patient], “I probably checked ZWIP 2 or 3 times” [Respondent 6, Informal caregiver], and “At the start I checked ZWIP once or twice a week. Later, it decreased” [Respondent 16, Patient].

The number of providers that patients had contact with was also highly variable. Whereas some patients never logged in, only added informal caregivers to their network, or waited for others to communicate, other patients used ZWIP intensively to communicate in their network: “Only the family practitioner and district nurse, with the other health care providers I did not have any contact” [Respondent 14, Informal caregiver].

Other patients mentioned that providers did not respond to the invitations for their networks and that they got little response to their messages from providers. Interviewed ZWIP users and informal caregivers used only the communication platform: “I only used the communication tool” [Respondent 10, Informal caregiver].

Figure 1. Average page views per month of usage for the 4 identified usage clusters.



XSL•FO
RenderX

Table 1. Comparison of frail older patients who used and did not use ZWIP.

Characteristics		Total	Used ZWIP	Did not use ZWIP	*P* value (2-sided)
Female, n (%)		400	180 (63.2)	220 (66.7)	.36
Education, n (%)					.25
	≤Primary education	136	56 (21.7)	80 (28.7)	
	Secondary education	380	192 (74.4)	188 (67.4)	
	University / tertiary education	21	10 (3.9)	11 (3.9)	
Marital status, n (%)					.43
	Married	253	128 (44.4)	125 (38.1)	
	Divorced	29	11 (3.8)	18 (5.5)	
	Widow / widower / partner deceased	290	132 (45.8)	158 (48.2)	
	Unmarried	38	15 (5.2)	23 (7)	
	Long-term cohabitation, unmarried	6	2 (0.7)	4 (1.2)	
Living independently, n (%)		314	134 (48.2)	180 (55.4)	.08
Informal caregiver, n (%)		484	242 (84.0)	242 (74.0)	.003**
Informal caregiver lives together with the frail older person, n (%)		242	132 (56.2)	110(46.2)	.03*
Benefit expected from more coordinated care according to family practitioner, n (%)		228	102 (39.5)	126 (43.9)	.30
Number of other care providers than the family practitioner involved, n (%)					.007**
	No other care providers	75	22 (7.6)	53 (16)	
	1-3 other care provider	407	200 (69.4)	207 (62.3)	
	More than 3 other care providers	138	66 (22.9)	72 (21.7)	
Age in yrs, mean (SD)		615	82.13 (5.77)	82.07 (5.72)	.89
Frailty index (scale 0-1, higher is worse), mean (SD)		616	0.29 (0.08)	0.30 (0.08)	.16
Barthel index (0-20, 20 completely independent), mean (SD)		616	16.01 (3.91)	15.99 (3.66)	.95
Morbidity, n (%)					
	Diabetes	146	68 (23.9)	78 (23.6)	.95
	Stroke	91	38 (13.3)	53 (16.1)	.34
	Heart failure	55	26 (9.1)	29 (8.8)	.88
	Cancer	76	39 (13.7)	37 (11.2)	.35
	COPD	100	44 (15.4)	56 (17)	.61
	Urinary incontinence	144	62 (21.8)	82 (24.8)	.37
	Osteoarthritis	2	1 (0.4)	1 (0.3)	.92
	Osteoporosis	39	16 (5.6)	23 (7)	.49
	Hip fracture	17	6 (2.1)	11 (3.3)	.35
	Fractures other than hip	12	7 (2.5)	5 (1.5)	.40
	Dizziness with falling	15	7 (2.5)	8 (2.4)	.98
	Benign prostate enlargement	16	10 (3.5)	6 (1.8)	.19
	Depression	10	2 (0.7)	8 (2.4)	.09
	Anxiety / panic disorder	16	11 (3.9)	5 (1.5)	.07
	Dementia	25	19 (6.7)	6 (1.8)	.002**

XSL•FO
RenderX

Characteristics		Total	Used ZWIP	Did not use ZWIP	*P* value (2-sided)
	Hearing problems	197	98 (34.4)	99 (30)	.25
	Vision problems	133	60 (21.1)	73 (22.1)	.75
Multimorbidity, mean (SD)		615	1.76 (1.42)	1.80 (1.34)	.68
Katz IADL (0-8, 8 completely limited), mean (SD)		616	5.06 (2.92)	5.46 (2.77)	.08
SF-36 Social activity limitations (1-5,with 5 no limitations), mean (SD)		616	1.58 (1.39)	1.57 (1.40)	.90
SF-36 mental health (1-100, 100 perfect mental health), mean (SD)		604	63.31 (13.11)	63.47 (13.12)	.90
Patient experience with being informed (1-5, with 5 optimal), mean (SD)		512	4.79 (0.67)	4.74 (0.78)	.43
Patient experience with coordination of care (1-5, with 5 optimal), mean (SD)		327	4.65 (0.88)	4.72 (0.82)	.45
Patient experience with co-decision making (1-5, with 5 optimal), mean (SD)		486	4.86 (0.54)	4.84 (0.60)	.71
Patient preferences for influence (1-5, with 5 optimal), mean (SD)		551	3.46 (1.13)	3.50 (1.10)	.68
Patient knowledge of providers (health and social) (1-5, with 5 optimal), mean (SD)		587	3.56 (0.65)	3.60 (0.69)	.38
Patient experience with self-management (1-5, with 5 optimal), mean (SD)		497	4.86 (0.52)	4.91 (0.45)	.20
Number of emergency family practitioner visits during past 12 months, mean (SD)		520	0.74 (2.08)	0.61 (1.22)	.38
Unplanned hospitalization during past 12 months, mean (SD)		535	3.52 (7.18)	3.85 (12.37)	.70
Degree of certainty about treatment (according to family practitioner) (0-10, with 10 complete certainty), mean (SD)		621	6.13 (2.42)	6.44 (2.38)	.10
Agreement between providers involved with the patient (according to family practitioner) coordination of care around patient (0-10, with 10 lot of experience), mean (SD)		621	5.68 (2.58)	6.02 (2.38)	.09

*$P<.05$.
**$P<.01$.

Table 2. Usage of the different functionalities of ZWIP, as used in the cluster analysis for the four different groups of users.

	Total group	Non-users	Regular users	Active users	Very active users
People using ZWIP, n (%)	290	211 (72.8)	67 (23.1)	9 (3.1)	3 (1.0)
Older people logged in at least once, n (%)	220	141 (64.1)	67 (30.5)	9 (4.1)	3 (1.4)
Number of providers in network per older person at baseline, mean (SD)	4.44 (2.77)	3.80 (2.20)	5.76 (3.28)	7.56 (3.40)	10.67 (0.58)
Page views per patient during 2 yrs, mean (SD)	102.64 (273.16)	19.32 (22.71)	188.22 (78.74)	659.33 (174.88)	2381.33 (490.02)
Messages sent in a network during 2 yrs, mean (SD)	11.54 (35.26)	2.47 (4.44)	19.94 (21.02)	70.89 (49.08)	284.33 (122.10)
Other ZWIP functions (health promotion usage, setting and modifying goals, actions to reach the goals) used during 2 yrs, mean (SD)	7.22 (6.62)	5.53 (3.23)	10.07 (7.86)	19.33 (12.83)	26.00 (27.78)

XSL•FO
RenderX

Textbox 1. Example user profiles per group.

Non-active user

A man in his 70s, suffering from heart failure and stroke. His daughter already predicted the limited use of ZWIP by her father, because his computer skills were poor. She lived very close to her father but also hardly used ZWIP for her father. Even though ZWIP seemed appealing to her, she logged in only three times. She agreed to participate in ZWIP in view of the future digitalization of society. The daughter found ZWIP redundant. The family practice and the pharmacy were across the street from where she lived, which facilitates face-to-face instead of digital consultation.

Regular user

A woman in her 90s, suffering from dementia, heart failure, and deafness. She was living together with her son, who managed her ZWIP. The family practitioner and the practice nurse participated in the patient's ZWIP network. Her son liked to occasionally use the ZWIP to ask questions regarding his mother's treatment and background information on her medical conditions. Her son also rescheduled his mother's appointments through the ZWIP. When his mother became more apractic due to dementia, the occupational therapist was invited to join ZWIP. Her son appreciated ZWIP for its user friendliness, the easy small-scale communication, and the complete overview of his mother's conditions and medication use.

Active user

A woman in her 80s, suffering from coronary heart disease. Her deceased husband was a health care provider. She managed ZWIP herself. Generally, she found it difficult to reach the family practice. What she appreciated most from ZWIP was that the family practitioner was easily accessible for health- or medication-related questions. That made her feel safe during the time of her myocardial infarction and it prevented, in her view, unnecessary face-to-face consultations in the family practice. She also kept her family practitioner informed of her condition after visits to the cardiologist. The family practitioner was her only active contact in ZWIP and that was all she needed. At the time of the interview, she was doing relatively well, so she hardly used ZWIP in that period.

Very active user

A man in his 80s, suffering from COPD and heart failure. He was highly educated and had good computer skills. He used ZWIP daily. He strongly appreciated the two main components of ZWIP: the messaging system and especially the goal-setting function. His family practitioner, cardiologist, and pulmonary nurse participated in his ZWIP. He reported that the frequency of the home visits has decreased and the physically exhausting face-to-face consultations with his health care providers had become rarer.

Perceptions of ZWIP Usage

Perceptions of ZWIP usage were rather positive, illustrated by the quotes below. Some patients and informal caregivers perceived that communication was quicker than by telephone and that ZWIP was a good way to co-decide on their own health and welfare issues with the provider: "It is fast, multidisciplinary and you are in control of your own care. Very pleasant!" [Respondent 22, Patient] and "You send a message, and when it is convenient for them they answer. I think it is easier for them, also for the family practitioner" [Respondent 20, Informal caregiver].

Exchange of information and access to providers was perceived to be better using ZWIP: "I think it (ZWIP) causes unnecessary appointments to be avoided and at the same time it provides a safe connection to the family practitioner" [Respondent 17, Patient].

Informal caregivers said the ZWIP gave them a complete overview of the patient's situation: "You can paint a more complete picture than on the phone." [Respondent 10, Informal caregiver] and "My parents needed more care. It seemed like a good idea to get a good overview and simplify communication" [Respondent 10, Informal caregiver].

Other participants were disappointed by the frequency of communication through ZWIP: "Well, we expected that there would be more information exchange..." [Respondent 20, Informal caregiver].

Predictors for Use: Barriers and Facilitators

In quantitative user analysis, several determinants of use were identified (Table 3). When a person had larger networks involved in their care according to the family physician, this led to more page views. Having an informal caregiver also contributed positively to ZWIP usage. Also, patients with a higher Barthel index (better ADL) at baseline showed fewer page views compared to the older people with lower Barthel index (worse ADL) at baseline. All patients slightly decreased in their use of ZWIP over time. On a family practice level, having a higher percentage of frail older people in the practice, previously having used electronic consultation methods, and being located in a village (versus a city) led to more active ZWIP usage.

Also, a major theme addressed in the interviews was facilitators of use. According to the interviewees, a successful older ZWIP user was interested in communicating about health and welfare status and was pro-actively sending messages to providers using ZWIP: "I like to be in control of my own care. I can handle different opinions. And I have several health problems, on which I have questions" [Respondent 22, Patient, on why ZWIP was successful in his case].)

Importantly, these patients had current, chronic health problems for which they used ZWIP. Patients using ZWIP should have computer skills and should have no cognitive problems, or should have an adequate informal caregiver who can use ZWIP on their behalf: "I can handle it (computer) very well. My mother cannot, so I support her" [Respondent 18, Informal

XSL•FO
RenderX

caregiver]. ZWIP was considered a good, quick, and easy communication tool: "I can easily communicate without picking up the phone for everything" [Respondent 10, Informal caregiver].

Informal caregivers, especially children who were living far away, mentioned that ZWIP made it easy for them to be involved in the care for their vulnerable parent. "It is a triangle, it is about my husband, but it could also support me" [Respondent 19, Informal caregiver] and "Before, you had to wait until the family practitioner came. Now we used ZWIP and the family practitioner was already informed when he came" [Respondent 3, Informal caregiver].

An important advantage compared to telephone and face-to-face communication was that in ZWIP it is possible to communicate at a time that is convenient for patients and providers: "A small example is that I do not have to take my mother to the family practitioner for everything. I can just ask a question" [Respondent 18, Informal caregiver].

Some patients mentioned that they used ZWIP because it gave them a sense of safety, being able to communicate with their health care providers and to ask questions in ZWIP: "It is good that I can alarm providers about my father's situation, rather than them hearing only his story" [Respondent 18, Informal caregiver]. One patient said that communication in ZWIP could replace a consultation, which was convenient for him because his medical condition made it difficult for him to visit his physicians: "Before, you had to wait until the family practitioner came. Now we used ZWIP and the family practitioner was already informed when he came" [Respondent 3, Informal caregiver].

Another major theme in the interviews was barriers for use. Patients or informal caregivers who did not or hardly used ZWIP often stated that there was no need for ZWIP, since they or the older people they cared for were healthy: "There was no reason for use. My father is actually still very healthy" [Respondent 15, Informal caregiver] and "Why should I communicate if I do not have any questions?" [Respondent 23, Patient].

Some older people simply had problems logging in but never asked for help, or they reported that support was not sufficient: "I thought it was all very complicated…I could not handle it" [Respondent 9, Informal caregiver] and "A computer is a little impossible for an older person" [Respondent 15, Informal caregiver].

However, patients did expect an increased usage of such interventions in the future when computer skills are more common among older patients. Other patients did not like the fact that ZWIP was a computer app and were not interested in using it. Some patients feared that the ZWIP would replace face-to-face contact with their health care providers and preferred to phone or visit the family practice: "You do not see your family practitioner anymore. That is a disadvantage… When contact is only digital, I think care becomes rather meager" [Respondent 16, Patient].

Other patients and informal caregivers would have wanted to use ZWIP but complained that they got no response or only a limited response from health care providers. This was particularly problematic if it was their family practitioner not responding. "We experienced slow reactions from health care providers…So we could not use it for communication" [Respondent 20, Informal caregiver].

Some of these professionals did not even accept the invitation to participate in the patient's network. Several patients said that they assumed that their family practitioners were too busy to communicate in ZWIP and so did not try further: "I found it cumbersome. And because I thought: they are never going to make time for that. They are already so busy" [Respondent 6, Informal caregiver].

XSL•FO
RenderX

Table 3. Association of potential patient-level and family practice/practitioner-level predictors with the number of total page views at baseline and across 2 years of follow-up using hierarchical linear models.

Baseline determinants (fixed effects)	Time-determinant determinants (linear rate of change)	Estimate	95% CI	*P* value
Intercept		-3.16	-19.41 to 13.09	.70
Patient-level determinants				
Network size (number of professionals at baseline)		2.75	1.83-3.67	<.001***
Education		1.39	-0.20 to 2.98	.09
Barthel index of ADL (0-20)		-0.94	-1.67 to -0.21	.012*
SF-36 mental health (0-100)		-0.15	-0.31 to 0.01	.07
Age at baseline		0.08	-0.29 to 0.45	.69
Multimorbidity		-0.75	-2.30 to 0.80	.35
Informal caregiver		6.67	0.71-12.63	.03*
Sex (female)		-3.32	-8.14 to 1.50	.18
Account closed		5.45	-10.76 to 21.66	.51
	Time (months from baseline)	-0.03	-1.48 to 1.42	.96
	time*network size (number of professionals at baseline)	-0.18	-0.26 to -0.10	<.001***
	time*education	-0.10	-0.28 to 0.08	.25
	time*barthel index at baseline	0.09	0.01-0.17	.02*
	time* SF-36 mental health at baseline	0.02	0.00-0.04	.06
	time*age at baseline	-0.05	-0.09 to -0.01	.03*
	time*multimorbidity at baseline	0.05	-0.13 to 0.23	.53
	time*informal caregiver at baseline	-0.55	-1.20 to 0.10	.10
	time*sex	0.43	-0.08 to 0.94	.11
	time*account closed	-0.20	-1.32 to 0.92	.72
Family practice-level determinants				
Village practice		24.25	15.06-33.44	<.001***
Electronic consultation		10.92	1.00-20.84	.03*
Percent frail		0.95	0.48-1.42	<.001***
	time*village	-0.16	-1.14-0.82	.75
	time*electronic consultation	-0.33	-1.39 to 0.73	.54
	time*percent frail	0.00	-0.06 to 0.06	.92

*P<.05.
**P<.01.
***P<.001.

Discussion

Principal Findings

Active use of ZWIP as an innovative POHC, specifically developed for frail older people, was limited, despite a small group of very active users. However, in active networks ZWIP was highly appreciated. ZWIP was mainly used for communication. The combination of quantitative and qualitative data revealed a distinct profile of an active ZWIP user: a patient or informal caregiver interested in communicating about their own or their relative's health and welfare status and sufficiently pro-active to actually stimulate communication in ZWIP. Importantly, these patients had current, chronic health problems that they used ZWIP for and had a lower functional performance (low Barthel scores). Multiple health care providers in a patient's network (reflected in network size) was also associated with increased ZWIP use.

Characteristics of Likely Users

The current study contributes to identifying characteristics of the older people who are likely to use POHCs—either on their own or with additional support—and which older people are unlikely to use a POHC. It seems that additional targeting criteria are useful for the successful implementation of POHCs for frail older people. Whereas some general characteristics

XSL•FO
RenderX

may be applicable for all potential users of eHealth interventions, others seem specific for this population of frail older people. In line with previous research [36], impaired health status (reflected in ADL functioning and indirectly in network size) was an important predictor for using a website in older people. This indicates that a minimal level of care need was required. So there needs to be a current reason for intensive contact with health care providers for a POHC to be used. However, other characteristics identified by previous research [37] such as education were not significantly related to ZWIP use. This may be due to the generally low level of education of the older people in the sample compared to younger populations. Additionally, patient age and gender seem to be a common predictor of usage [33]; however, in this study neither was significant. This may be due to the low level of ZWIP's use. At the same time, the effect of age is partially captured by the health status variable, as shown by a small and significant negative interaction between age and continued usage (analysis on request). Further research is needed to investigate how age influences usage beyond health status.

Two additional characteristics of successful implementation are worth mentioning. On a family practice level, this study shows that a high percentage of frail older people in the family practice's patient population led to more ZWIP use. This finding suggests that context effects are also important for usage. Additionally, having an informal caregiver significantly contributed to ZWIP's increased usage, which may indicate that informal caregivers can overcome problems of (generally) low computer literacy among older people, as suggested by previous research [38,39].

According to interviewees, future active ZWIP users are interested in managing their own (or their loved one's) care and experiencing current health problems. Such observations are consistent with major dimensions of the theoretical models of information and communication technology adoption, such as the Unified Theory of Acceptance and Use of Technology model, namely performance expectation, effort expectancy, facilitating conditions, computer experience and voluntariness of use, and social influence [33]. Performance expectancy is defined as the degree to which individuals believe that using the system will help them attain gains in performance. Effort expectancy is the degree of ease associated with the use of the system. Facilitating conditions are defined as the degree to which an individual believes that an organizational and technical infrastructure exists to support use of the system. Computer experience is if patients have used computers before, voluntariness of use is how far patients felt they had a choice in using the system, and social influence is having the feeling that others also use the system and their opinions on the system. Specifically, in terms of performance expectations, patients were more likely to use ZWIP if they expected improved communication with the physician. In terms of effort expectancy, on average they perceived ZWIP use to be easy, although there is a group who find the computer a real barrier to ZWIP use. Patients were also more likely to use ZWIP if they were facilitated in ZWIP usage or had previous computer experience.

Additionally, ZWIP usage was more likely if patients voluntarily participated in ZWIP, and in terms of social influence, actually got messages. Therefore, it seems that the successful use of patient networks requires providers to be willing to respond to patient messages and actively send messages to patients. Conversely, health care providers who decline invitations or fail to respond to patient messages form a serious barrier for implementation. Further research is needed to quantitatively explore the role of technology acceptance factors, for example, using the Unified Theory of Acceptance and Use of Technology model [33] to guide implementation of POHCs in older populations. However, this first study to identify the characteristics of successful ZWIP participants already allows for a stricter selection of patients who might most easily benefit from such interventions. As long as a large group of frail persons experiences the electronic nature of POHCs as a major barrier to use, the efficacy and effectiveness of POHCs can be better tested in more selective samples [40,41]. Carefully selected older people should be more successful in digital self-management.

Implications for Implementation and Practice

Although we have found some examples of frail older patients in their 80s successfully using ZWIP, low computer literacy of older people seems to hamper the implementation efforts. In the Netherlands, only 39% of older people aged 75 years and older report having Internet access [42], whereas in the younger populations this is close to 100%. One option is to give older people computer training. Previous research has shown that training older people to use computers and health websites that they are expected to use can be successful [43]. Therefore, we have attempted train older people in using computers and the ZWIP website [44]. However, older people were not receptive to our chosen method of implementation of receiving trainers in their homes [44]. An alternative is involving the informal caregivers at an earlier stage, as both our qualitative results and quantitative results revealed the importance of having an informal caregiver for continued use. It is likely that the general level of computer literacy in older people will increase over time. However, low computer literacy will remain problematic in disadvantaged groups or very old people for a considerable time and may be a persistent problem for older people suffering cognitive impairment. In future studies, various implementation methods should be compared to raise computer literacy in the older people and to increase the willingness of informal caregivers to participate. In addition, the effect of using eHealth interventions on various components of health—namely physical, social, and mental health—should continue to be explored further.

Strengths and Limitations

This study reports a process evaluation of POHC usage in a population of frail older people. Its major strength lies in the mixed methods approach, combining both qualitative and quantitative information on usage, and in its relatively large sample size. Its outcome, a profile of a successful older eHealth user, makes it possible to target likely users. Adequately targeting audiences who will use a new intervention, represents a precondition for studies aiming to show the benefits of such

XSL•FO
RenderX

interventions. Thus, these profiles form a valuable contribution to future research and implementation projects of eHealth interventions.

The study also has some limitations to consider. First, this process evaluation of usage was conducted in the Netherlands within a strong primary care system and so it can be generalized only to countries with a similar system. Additionally, we did not continuously measure the health status of individuals, which was shown qualitatively as a main driver of usage, and this should be explored in further research. Finally, while we explored patients' levels of computer literacy, access to computers, and user-friendliness of ZWIP for the individual older person in the interviews, we omitted a quantitative assessment of these issues and this should be pursued further during evaluation of other applications.

Conclusions

The current evaluation of use of the ZWIP, a personal online health community for frail older people, revealed important predictors of usage of eHealth interventions in older populations. Frail older patients with poor health status reflected in functional problems, with at least 5-6 providers involved perceived the most benefit and therefore used ZWIP more actively. Sufficient computer skills in either the patient or the informal caregiver and an interest to play an active role in their own care were essential. Therefore, during implementation of POHCs, and probably other eHealth interventions, informal caregivers should be involved from the start of the project. Profiling successful users can facilitate more effective targeting of frail older patients for implementation of eHealth interventions in the near future.

Acknowledgments

We would like to thank Wilma Derksen-Driessen, Emile ter Horst Mark Kuster, Charlotte Neger, Jean Nielen, Joep Scheltinga, and Leontien van Nieuwenhuijzen for their assistance in the development and implementation of the program. This study was supported by the Netherlands Organisation for Health Research and Development (ZonMw) project number 311050201. The funder had no role in conducting the study, analysis of the data, or during the publication of the manuscript.

Authors' Contributions

The authors contributed to the manuscript as follows: PM drafted the manuscript; HS, MP, SHMR contributed to acquisition of data; PM did the statistical analysis of data; PM, MP, SHMR, MGMOR, RJFM interpreted the data. All authors contributed to the study design, revised the manuscript for important intellectual content, and approved the manuscript. MGMOR and RJFM supervised the study and obtained funding.

Conflicts of Interest

None declared.

Multimedia Appendix 1

ZWIP usage.

[WMV File (Windows Media Video), 15MB - jmir_v16i12e278_app1.wmv]

Multimedia Appendix 2

Interview questions.

[PDF File (Adobe PDF File), 4KB - jmir_v16i12e278_app2.pdf]

References

1. Rockwood K, Mitnitski A. Frailty in relation to the accumulation of deficits. J Gerontol A Biol Sci Med Sci 2007 Jul;62(7):722-727. [Medline: 17634318]
2. Robben SH, Huisjes M, van Achterberg T, Zuidema SU, Olde Rikkert MG, Schers HJ, ZOWEL NN Study Group. Filling the Gaps in a Fragmented Health Care System: Development of the Health and Welfare Information Portal (ZWIP). JMIR Res Protoc 2012;1(2):e10 [FREE Full text] [doi: 10.2196/resprot.1945] [Medline: 23611877]
3. Vedel I, Akhlaghpour S, Vaghefi I, Bergman H, Lapointe L. Health information technologies in geriatrics and gerontology: a mixed systematic review. J Am Med Inform Assoc 2013;20(6):1109-1119 [FREE Full text] [doi: 10.1136/amiajnl-2013-001705] [Medline: 23666776]
4. van den Berg N, Schumann M, Kraft K, Hoffmann W. Telemedicine and telecare for older patients--a systematic review. Maturitas 2012 Oct;73(2):94-114. [doi: 10.1016/j.maturitas.2012.06.010] [Medline: 22809497]
5. van der Eijk M, Faber MJ, Aarts JW, Kremer JA, Munneke M, Bloem BR. Using online health communities to deliver patient-centered care to people with chronic conditions. J Med Internet Res 2013;15(6):e115 [FREE Full text] [doi: 10.2196/jmir.2476] [Medline: 23803284]

XSL•FO
RenderX

6. Demiris G. The diffusion of virtual communities in health care: concepts and challenges. Patient Educ Couns 2006 Aug;62(2):178-188. [doi: 10.1016/j.pec.2005.10.003] [Medline: 16406472]
7. Makai P, Perry M, Robben S, Schers HJ, Heinen MM, Olde Rikkert MG, et al. Evaluation of an eHealth intervention in chronic care for frail older people: why adherence is the first target. J Med Internet Res 2014;16(6):e156 [FREE Full text] [doi: 10.2196/jmir.3057] [Medline: 24966146]
8. Kim EH, Stolyar A, Lober WB, Herbaugh AL, Shinstrom SE, Zierler BK, et al. Challenges to using an electronic personal health record by a low-income elderly population. J Med Internet Res 2009;11(4):e44 [FREE Full text] [doi: 10.2196/jmir.1256] [Medline: 19861298]
9. DesRoches CM, Campbell EG, Rao SR, Donelan K, Ferris TG, Jha A, et al. Electronic health records in ambulatory care--a national survey of physicians. N Engl J Med 2008 Jul 3;359(1):50-60. [doi: 10.1056/NEJMsa0802005] [Medline: 18565855]
10. Jha AK, Doolan D, Grandt D, Scott T, Bates DW. The use of health information technology in seven nations. Int J Med Inform 2008 Dec;77(12):848-854. [doi: 10.1016/j.ijmedinf.2008.06.007] [Medline: 18657471]
11. Jha AK, DesRoches CM, Campbell EG, Donelan K, Rao SR, Ferris TG, et al. Use of electronic health records in U.S. hospitals. N Engl J Med 2009 Apr 16;360(16):1628-1638. [doi: 10.1056/NEJMsa0900592] [Medline: 19321858]
12. Jha AK, Bates DW, Jenter C, Orav EJ, Zheng J, Cleary P, et al. Electronic health records: use, barriers and satisfaction among physicians who care for black and Hispanic patients. J Eval Clin Pract 2009 Feb;15(1):158-163. [doi: 10.1111/j.1365-2753.2008.00975.x] [Medline: 18759752]
13. Poon EG, Blumenthal D, Jaggi T, Honour MM, Bates DW, Kaushal R. Overcoming barriers to adopting and implementing computerized physician order entry systems in U.S. hospitals. Health Aff (Millwood) 2004;23(4):184-190 [FREE Full text] [Medline: 15318579]
14. Mair FS, May C, O'Donnell C, Finch T, Sullivan F, Murray E. Factors that promote or inhibit the implementation of e-health systems: an explanatory systematic review. Bull World Health Organ 2012 May 1;90(5):357-364 [FREE Full text] [doi: 10.2471/BLT.11.099424] [Medline: 22589569]
15. Verhoeven F, van Gemert-Pijnen L, Dijkstra K, Nijland N, Seydel E, Steehouder M. The contribution of teleconsultation and videoconferencing to diabetes care: a systematic literature review. J Med Internet Res 2007;9(5):e37 [FREE Full text] [doi: 10.2196/jmir.9.5.e37] [Medline: 18093904]
16. Paré G, Moqadem K, Pineau G, St-Hilaire C. Clinical effects of home telemonitoring in the context of diabetes, asthma, heart failure and hypertension: a systematic review. J Med Internet Res 2010;12(2):e21 [FREE Full text] [doi: 10.2196/jmir.1357] [Medline: 20554500]
17. Cartwright M, Hirani SP, Rixon L, Beynon M, Doll H, Bower P, Whole Systems Demonstrator Evaluation Team. Effect of telehealth on quality of life and psychological outcomes over 12 months (Whole Systems Demonstrator telehealth questionnaire study): nested study of patient reported outcomes in a pragmatic, cluster randomised controlled trial. BMJ 2013;346:f653 [FREE Full text] [Medline: 23444424]
18. van Kempen JA, Schers HJ, Jacobs A, Zuidema SU, Ruikes F, Robben SH, et al. Development of an instrument for the identification of frail older people as a target population for integrated care. Br J Gen Pract 2013 Mar;63(608):e225-e231 [FREE Full text] [doi: 10.3399/bjgp13X664289] [Medline: 23561790]
19. Olde-Rikkert MG, Long JF, Philp I. Development and evidence base of a new efficient assessment instrument for international use by nurses in community settings with older people. Int J Nurs Stud 2013 Sep;50(9):1180-1183. [doi: 10.1016/j.ijnurstu.2012.08.007] [Medline: 22980483]
20. Extermann M, Aapro M, Bernabei R, Cohen HJ, Droz JP, Lichtman S, Task Force on CGA of the International Society of Geriatric Oncology. Use of comprehensive geriatric assessment in older cancer patients: recommendations from the task force on CGA of the International Society of Geriatric Oncology (SIOG). Crit Rev Oncol Hematol 2005 Sep;55(3):241-252. [doi: 10.1016/j.critrevonc.2005.06.003] [Medline: 16084735]
21. Creswell JW, Klassen A, Plano Clark V, Smith KC. National Institutes of Health. Bethesda (Maryland); 2011. Best practices for mixed methods research in the health sciences URL: http://obssr.od.nih.gov/mixed_methods_research/pdf/Best_Practices_for_Mixed_Methods_Research.pdf [accessed 2014-11-18] [WebCite Cache ID 6UB84wiAq]
22. Lingard L, Albert M, Levinson W. Grounded theory, mixed methods, and action research. BMJ 2008;337:a567. [Medline: 18687728]
23. Wade DT, Collin C. The Barthel ADL Index: a standard measure of physical disability? Int Disabil Stud 1988;10(2):64-67. [Medline: 3042746]
24. Collin C, Wade DT, Davies S, Horne V. The Barthel ADL Index: a reliability study. Int Disabil Stud 1988;10(2):61-63. [Medline: 3403500]
25. Katz S. Assessing self-maintenance: activities of daily living, mobility, and instrumental activities of daily living. J Am Geriatr Soc 1983 Dec;31(12):721-727. [Medline: 6418786]
26. 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]
27. Beckstead JW. Using hierarchical cluster analysis in nursing research. West J Nurs Res 2002 Apr;24(3):307-319. [Medline: 11911382]

XSL•FO
RenderX

28. Ward JH. Hierarchical Grouping to Optimize an Objective Function. Journal of the American Statistical Association 1963 Mar;58(301):236-244. [doi: 10.1080/01621459.1963.10500845]
29. Calinski T, Harabasz J. A dendrite method for cluster analysis. Comm. in Stats. - Theory & Methods 1974;3(1):1-27. [doi: 10.1080/03610927408827101]
30. Singer JD. Using SAS PROC MIXED to Fit Multilevel Models, Hierarchical Models, and Individual Growth Models. Journal of Educational and Behavioral Statistics 1998 Jan 01;23(4):323-355. [doi: 10.3102/10769986023004323]
31. Barbour RS. Checklists for improving rigour in qualitative research: a case of the tail wagging the dog? BMJ 2001 May 5;322(7294):1115-1117 [FREE Full text] [Medline: 11337448]
32. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
33. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: Toward a unified view. MIS quarterly 2003:425.
34. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
35. Creswell JW, Fetters MD, Ivankova NV. Designing a mixed methods study in primary care. Ann Fam Med 2004;2(1):7-12 [FREE Full text] [Medline: 15053277]
36. Chou WY, Hunt YM, Beckjord EB, Moser RP, Hesse BW. Social media use in the United States: implications for health communication. J Med Internet Res 2009;11(4):e48 [FREE Full text] [doi: 10.2196/jmir.1249] [Medline: 19945947]
37. Kakkar A, Jacobson BC. Failure of an Internet-based health care intervention for colonoscopy preparation: a caveat for investigators. JAMA Intern Med 2013 Jul 22;173(14):1374-1376. [doi: 10.1001/jamainternmed.2013.6477] [Medline: 23877080]
38. Sadasivam RS, Kinney RL, Lemon SC, Shimada SL, Allison JJ, Houston TK. Internet health information seeking is a team sport: analysis of the Pew Internet Survey. Int J Med Inform 2013 Mar;82(3):193-200. [doi: 10.1016/j.ijmedinf.2012.09.008] [Medline: 23149121]
39. Mazor KM, Williams AE, Roblin DW, Gaglio B, Cutrona SL, Costanza ME, et al. Health literacy and pap testing in insured women. J Cancer Educ 2014 Dec;29(4):698-701. [doi: 10.1007/s13187-014-0629-7] [Medline: 24633725]
40. Bates DW, Wells S. Personal health records and health care utilization. JAMA 2012 Nov 21;308(19):2034-2036. [doi: 10.1001/jama.2012.68169] [Medline: 23168828]
41. Makai P, Melis RJ, Olde-Rikkert MG. Technical difficulties and evaluating e-health interventions. JAMA Intern Med 2014 Feb 1;174(2):304-305. [doi: 10.1001/jamainternmed.2013.12753] [Medline: 24493611]
42. Statistics Netherlands. Rapid increase of internet usage among older people. 2013. URL: http://www.cbs.nl/en-GB/menu/themas/vrije-tijd-cultuur/publicaties/artikelen/archief/2013/2013-4005-wm.htm [accessed 2014-11-18] [WebCite Cache ID 6UB8ghV5Z]
43. Xie B, Bugg JM. Public library computer training for older adults to access high-quality Internet health information. Libr Inf Sci Res 2009 Sep 1;31(3):155 [FREE Full text] [doi: 10.1016/j.lisr.2009.03.004] [Medline: 20161649]
44. Robben SH, Perry M, Huisjes M, van Nieuwenhuijzen L, Schers HJ, van Weel C, et al. Implementation of an innovative web-based conference table for community-dwelling frail older people, their informal caregivers and professionals: a process evaluation. BMC Health Serv Res 2012;12:251 [FREE Full text] [doi: 10.1186/1472-6963-12-251] [Medline: 22894654]

Abbreviations

ADL: Activities of Daily Living
Easycare-TOS: Two-step Older person Screening
IADL: Instrumental Activities of Daily Living
POHC: Personal Online Health Community
SF-36: Short Form-36 Health Survey
ZWIP: Zorg en Welzijn Informatie Portaal (Health and Welfare portal)

XSL•FO
RenderX

Edited by G Eysenbach; submitted 17.06.14; peer-reviewed by R Greysen, T Luger2; comments to author 07.08.14; revised version received 04.09.14; accepted 20.10.14; published 17.12.14.

Please cite as:
Makai P, Perry M, Robben SHM, Schers H, Heinen M, Olde Rikkert MGM, Melis RJF
Which Frail Older Patients Use Online Health Communities and Why? A Mixed Methods Process Evaluation of Use of the Health and Welfare Portal
J Med Internet Res 2014;16(12):e278
URL: http://www.jmir.org/2014/12/e278/
doi:10.2196/jmir.3609
PMID:25519769

©Peter Makai, Marieke Perry, Sarah HM Robben, Henk Schers, Maud Heinen, Marcel GM Olde Rikkert, René JF Melis. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 17.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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

Exploring the Relationship Between Changes in Weight and Utterances in an Online Weight Loss Forum: A Content and Correlational Analysis Study

Eric B Hekler[1], PhD; Gaurav Dubey[2], MS; David W McDonald[3], PhD; Erika S Poole[2], PhD; Victor Li[3], BA; Elizabeth Eikey[2], BS

[1]School of Nutrition and Health Promotion, Arizona State University, Phoenix, AZ, United States

[2]Pennsylvania State University, University Park, PA, United States

[3]University of Washington, Seattle, WA, United States

Corresponding Author:
Eric B Hekler, PhD
School of Nutrition and Health Promotion
Arizona State University
500 N. 3rd St.
Phoenix, AZ, 85003
United States
Phone: 1 602 827 2271
Fax: 1 602 827 2253
Email: ehekler@asu.edu

Abstract

Background: There is increasing interest in the use of online forums as a component of eHealth weight loss interventions. Although the research is mixed on the utility of online forums in general, results suggest that there is promise to this, particularly if the systems can be designed well to support healthful interactions that foster weight loss and continued engagement.

Objective: The purpose of this study was to examine the relationship between the styles of utterances individuals make on an online weight loss forum and week-to-week fluctuations in weight. This analysis was conducted to generate hypotheses on possible strategies that could be used to improve the overall design of online support groups to facilitate more healthful interactions.

Methods: A convenience sample of individuals using an online weight loss forum (N=4132) included data both on online forum use and weight check-in data. All interactions were coded utilizing the Linguistic Inquiry and Word Count (LIWC) system. Mixed model analyses were conducted to examine the relationship between these LIWC variables and weight over time.

Results: Results suggested that increased use of past-tense verbs (P=.05) and motion (P=.02) were associated with lower weekly weights whereas increased use of conjunctions (eg, and, but, whereas; P=.001) and exclusion words (eg, but, without, exclude; P=.07) were both associated with higher weight during the weeks when these utterances were used more.

Conclusions: These results provide some insights on the styles of interactions that appear to be associated with weight fluctuations. Future work should explore the stability of these findings and also explore possibilities for fostering these types of interactions more explicitly within online weight loss forums.

(J Med Internet Res 2014;16(12):e254) doi:10.2196/jmir.3735

KEYWORDS

social media; social support; weight loss; Natural Language Processing

Introduction

Background

The Internet is becoming a place for health information and social support. A recent Pew Research Center survey reported 72% of respondents search for health-related information online at least once a year, with more than a quarter of survey respondents specifically seeking experiences of others with similar health concerns [1]. Online communities for health not only offer social support [2-4], accountability [5], and peer modeling [6], but can also extend the reach of care for patients who fall between the cracks of traditional health care. Based on

XSL•FO
RenderX

this, there is great potential for online forums as a key part of systems that foster health behavior change.

Although there is great promise, current evidence supporting the utility of online communities is limited [7-9]. Specifically, these online communities, including social media and online forums/message boards, often promote only modest behavior change. This seems to be particularly the case because of high attrition rates. Based on this, there is great potential for these systems to promote behavior change and weight loss at scale, but additional research is required to better understand how best to design the systems to facilitate continued use and weight loss.

One strategy that could be utilized to improve the design of these systems is to explore how interactions within a system relate to outcomes such as weight. A great deal of research from psychology has articulated styles of interactions during therapeutic sessions that are predictive of behavior change. For example, empathic regard by a therapist and committal statements made by a client, particularly at the end of the session, are both predictive of improved clinical outcomes, such as reduced alcohol consumption [10-12]. Beyond therapeutic style, previous research by Pennebaker and colleagues [13,14] have provided a solid foundation for computationally documenting and understanding interaction styles from written documents (more subsequently). Collectively, this work suggests that there is likely great potential for understanding how best to improve online forums by articulating the interactions that appear to be associated with improved outcomes.

This paper reports findings concerning the relationship between styles of utterances in an online forum and weight loss via data provided to us from the makers of DropPounds (note: name changed to protect confidentiality). DropPounds is one of the most popular weight loss smartphone and Web-based online weight loss forums currently on the market. Users of DropPounds can specify various types of personal goals (eg, target weight, exercise, measurements, nutrition). DropPounds then prepares personalized targets to strive for and, in the paid premium plans, detailed plans. DropPounds provides a companion application for smartphones, tablets, and PCs. This application allows users to log their exercise and calorie consumption. The service also integrates with wireless activity and weight tracking devices, including the Withings scale, Nike+, and Fitbit. DropPounds provides peer support in the form of an online forum, along with social networking features such as private messaging and friend lists. The forum can be used to facilitate team and open challenges. It also provides a venue for discussing health, nutrition, exercise, personal reflection and experiences, and various other topics (eg, see [15]). In this paper, we examine how different utterances of participants in the DropPounds online forums, as codified by a validated natural language processing tool called Linguistic Inquiry and Word Count (LIWC), predict week-to-week fluctuations in weight among DropPounds forum users.

Related Research

Internet-Assisted Weight Loss Interventions

Although online forums are used for many purposes, research studying the effect of online forums as a tool to support weight loss is limited. The vast majority of research on online weight loss interventions (eg, Tate et al [16], Womble et al [17], Finley et al [18], Harvey-Berino et al [19], Neve et al [20,21], and Weinstein [22]) examines the effect of Internet-assisted weight loss intervention packages that included an online forum as 1 of its components. There is far less work examining characteristics of the forum itself that might predict health outcomes. Bensley et al [23] conducted one such study; they reviewed online groups and generated 5 possible design suggestions, including (1) recreate the human experience, (2) personalize, (3) create a dynamic experience, (4) provide a supportive environment, and (5) build on sound theory. This research was largely qualitative, however, with limited empirical data supporting these design suggestions. There is clearly potential for deeper research in this area. Our research attempts to fill this research gap by examining the effects of forum activity on weight loss.

Language Inquiry and Word Count

LIWC is a linguistic analysis tool [13,24-26]. It consists of a dictionary mapping words and word stems to linguistic categories, as well as a software program that calculates various metrics on text data using this dictionary. (Note that we did not use the program itself for this project. Because the amount of data we were working with was beyond the tool's abilities, we used a custom version of a Perl script.) The purpose of LIWC is "to provide a method for studying psychological phenomena (eg, emotion, cognition) that are present in individuals' communications" [24,25].

When LIWC processes a piece of text, counts are maintained corresponding to each category in its dictionary. When a word is scanned, the counts for all dictionary categories that include that word are incremented. The output consists of the values for all counts once the entire text has been scanned.

Dictionary

LIWC's dictionary assigns words and word stems to various categories (eg, positive emotion, anxiety, pronouns). The categories of the current dictionary can themselves be categorized in the following ways [13,24-26]: (1) 22 standard linguistic dimensions (eg, pronouns, articles, auxiliary verbs), (2) 32 categories tapping psychological constructs (eg, affect, cognition, biological processes), (3) 7 personal concern categories (eg, work, home, leisure activities), and (4) 3 paralinguistic dimensions (assents, fillers, nonfluencies).

LIWC's dictionary has gone through multiple iterations of compilation and revision [13]. Initially, there were only 2 categories—positive and negative emotion—but this was expanded during the original round of compilation as the compilers added more types of language they wanted to capture. Initial word lists were assembled from various sources (eg, dictionaries) and assessed by groups of independent judges. Initial compilation was done between 1992 and 1994, and revisions were done in 1997 and 2007 [13].

Linguistic Inquiry and Word Count Validation

LIWC has been validated through a variety of mechanisms. For example, in a writing study, college students were asked to write

XSL•FO
RenderX

essays from both emotional and nonemotional standpoints. These essays were rated using both LIWC and a panel of 4 independent judges [27]. Later analyses of these data found that the LIWC scales and judges' ratings were highly correlated [13,24-26].

There is a research precedent for using LIWC analysis of text to study various behavioral and psychological issues. For instance, Rude et al [28], Gunsch et al [29], and Kowalski [30] used LIWC analysis to gain insight into the cognitive operations associated with depression, the affective composition of political campaign ads, and the emotional perception of teasing, respectively.

On the other hand, the use of LIWC metrics as predictive tools has also been investigated. For instance, Simmons et al [31], Sexton and Helmreich [32], and Leshed et al [33] used LIWC metrics to predict marital health, pilot performance, and performance during collaborative tasks. We have applied this research methodology to the current task. DropPounds is ideal for this sort of analysis because natural writing is a fundamental part of its feature set and the sort of disclosures that occur within the forums map on to the intended purpose of LIWC focused on self-disclosure [27].

Research Questions

Overview

We explored 4 different models examining the relationship between utterance styles and weekly fluctuations in weight. Note that for each model we also included word count in all models as a control for overall verbal fluency.

Model 1

The first model focused on discussions about health. It is logical to presume that individuals who tend to talk more about health, consumption, or general biological mechanisms of weight loss may be more actively engaging in their weight loss during those times and, thus, lose more weight. Based on this presumption, we focused on content words as our first model.

Model 2

The second model focused on emotion. Specifically, it is plausible that content words related to positive and negative affect would be predictive of weight fluctuations with more positive affect associated with weight loss and more negative affect associated with weight gain.

Model 3

In our third model, we chose groupings of words that have been linked to psychological phenomena in previous research that could impact weight loss, including (1) social process words (eg, talk, they, child) which have been linked to social support; (2) past tense use (eg, went, ran, had), which suggests a focus on the past; (3) certainty (eg, always, never), which has been linked with emotional stability; (4) exclusive words (eg, but, without, exclude), which have been linked with honesty; and (5) third-person singular (eg, she, her, him), which has been linked to exhibiting social support [13,14].

Model 4

The fourth model was data-driven, in that we chose styles of utterances that were observed in our dataset most often within the forums and could be indicative of important areas to focus on within online weight loss forums. This fourth model specifically included past tense use, conjunctions, and "motion" words (eg, arrive, go, car).

Per Pennebaker et al [13,14], although a lot of initial emphasis was placed on "content" words (eg, emotions and personal concerns) in early development of LIWC, research increasingly indicated that function or "style words" (eg, pronouns, prepositions, verb tense) often had higher association with phenomena under study. As such, we anticipated the majority of findings to be localized around these style words, but we did develop content-related explorations based on their intuitive appeal and because there is limited research exploring these questions.

Methods

Data Preparation

Two datasets provided by the makers of DropPounds were combined to conduct these analyses. The first included the following data at a weekly level for users of DropPounds: weight, number of active days per week, calorie intake budget, number of calories consumed, and number of calories burned through exercise. In addition to users entering data manually, a subset of users also used Withings brand wireless-enabled weight scales. Although users may check in these data several times a week, when we received the data, it had been aggregated and reported on a weekly basis. Every check-in's weight was measured at the end of a weekly interval. This interval always started on the same day of the week—either Sunday or Monday, depending on the user. This check-in date was created based on the data provided to us (ie, the weekly weight check-in data was always aggregated on a Sunday or Monday). Identifying each user's check-in day was necessary for aggregating weekly LIWC metrics from the values for individual posts.

The second dataset included all forum posts of the participants. Because weight and other health data check-ins were aggregated weekly, we also aggregated forum-derived data at the weekly level using the same start weekday. For each post, we generated LIWC metrics.

After data cleaning, the dataset included weekly weight loss progress and forum check-ins from DropPounds users between July 2008 and October 2012. We processed the forum posts using the LIWC linguistic analysis tool, generating content metrics for each forum post based on a dictionary that assigns words of the English language to various topic categories as already discussed. These metrics were used as predictors in a multilevel mixed model analysis, with user weight as the predicted outcome variable.

Participants within the study were self-selected. Specifically, the act of joining the forum occurred via finding the forum online (either via being directed from DropPounds mobile app or likely via a Google Search). There was no other active

XSL•FO
RenderX

recruitment for this forum beyond standard commercial practices that might be deployed by a commercial weight loss company.

Linguistic Inquiry and Word Count Metrics

LIWC results were generated for every single forum post in the forum dataset and then aggregated to the weekly level to conform to the weight metric. Because LIWC metrics are word counts, aggregating each LIWC metric for a given user over a given week simply involved adding up the values of the metric for the user's posts over that week. Word count was controlled for in all models.

Data Cleaning

Data Anomalies

The weekly check-in data contained a number of anomalies, such as outliers, out-of-range check-ins (ie, automatically generated check-ins from before the earliest or after the latest possible check-in date), and sequences of check-ins that consisted of the same weight carried forward from previous weeks. We applied the following filters and corrections to clean the dataset.

1. If 2 or more consecutive check-ins had diet and exercise check-in values of zero (an indicator of nonreporting) and they had identical weight, then we assumed that those consecutive weight check-ins were automatically generated and filtered them out of our dataset.
2. If a check-in had nonzero diet and exercise check-in data, and was followed by a check-in with identical weight and diet and exercise check-in data that equaled zero, we assumed the second check-in was automatically generated by copying the first and we filtered out the second check-in.
3. We flagged all users with weight check-ins over 600 lb for outlier investigation.
4. We examined the flagged data and corrected anomalous weight entries where possible, based on surrounding entries (ie, a string of entries with weight 179 lb contained a single check-in with weight 1111 lb and this was corrected by inspection).
5. We used linear interpolation to correct sequences of repeated outlier check-ins.
6. We deleted all users who showed no change in weight throughout their data-gathering period.
7. Two users had parallel data streams. We deleted the automatically generated streams, leaving the manually checked-in data.
8. We calculated the change in weight from 1 week to the next for each check-in. If there was a large weight change between check-ins, we took the length of time gap was taken into account before applying outlier processing. To filter out users with unreasonable changes in weight, we only kept check-ins for users who had a maximum absolute weekly weight change of 5 pounds or less, which is conceptually similar to strategies used by Finley et al [18].

In applying these changes, we brought the size of the dataset from more than 1 million check-ins to approximately 100,000, representing approximately 4000 users. However, further data preparation was necessary before we could begin modeling. We describe this process in the following section.

To provide some context for our statistical results, we included text from the forums in the Post Text Results section within the current paper. Note that slight modifications to the posts were made if any potentially identifying information was provided in the text. To find these posts, we selected the top 10 posts for each category (ie, high in the particular style word use, such as past-tense verbs). We then selected passages from this group that were the best demonstration of the general style of conversation found among those 10 posts.

Statistical Analyses

Similar to Neve et al [20], we chose multilevel mixed models as our analysis technique [34]. The advantage of this technique is that it is capable of dealing with observations that are nested (eg, multiple observations for each individual), have unbalanced data (where individuals have unequal observations), and contain missing values.

Prior to conducting any mixed model analyses, variables are recentered to make the results more readily interpretable. A recentered variable is restated as a deviation from a theoretically meaningful reference value (eg, time since first observation instead of calendar date and time).

In our dataset, we recentered the starting date for each week. This was done by representing starting dates as week numbers. Each starting date was replaced by the number of weeks between the starting date and the date when the user first used DropPounds.

Multilevel mixed models were developed using an iterative methodology. Each model attempted to improve on the previous model, reducing the residual error and explaining more of the target variance. Here, we describe the models generated.

The first model generated was an unconditional means model. This model included no predictors and attempted to express the target variable in terms of a grand mean (the fixed intercept) and a user-specific deviation from the grand mean (the random intercept). This model served as a baseline for model comparison. If the random effect was not statistically significant, we concluded that the individual-level mean target values were not statistically different than the grand mean target value for the entire dataset.

The next set of models added time as a predictor. These were the unconditional growth models and they attempted to determine the effect of time on the target variable. Multiple unconditional growth models were developed. Each included a higher-order power of time than the previous. This enabled us to study curvilinear change over time, if there was a plausible reason to anticipate such change (as was the case with weight whereby it was plausible for a quartic linear effect based on the well-studied concept of "yo-yo" dieting with weight loss, plateaus, and then weight gain again and another plateau). Finally, we generated the models that include the predictors of interest after controlling for time and for overall word count.

XSL•FO
RenderX

Results

Mixed Model Results

Table 1 reports descriptive statistics about the sample (N=4132). In brief, the mean age was 37.5 (SD 12.4) years and most were women (81.60%, 3371/4132). These values are similar to those seen in other studies focused on commercial weight loss interventions [18,20].

Table 1. Descriptive statistics of the sample (N=4132).[a]

Variables	Sample	Range
Age (years), mean (SD)	37.5 (12.4)	18-87
Starting weight (lbs), mean (SD)	190.1 (46.7)	63-483
Number of weeks actively checked-in weight and used forum, mean (SD)	23.5 (23.6)	2-185
Women, n (%)	3371 (81.60)	

[a] Because these data were provided to us, "normal" demographic data for scientific publications (eg, ethnicity, BMI) were not available.

Results comparing the unconditional means model to the unconditional growth models indicated that there was a significant impact of time on weight. Specifically, results indicated that there was a significant effect of linear and quadratic time. These results suggested that individuals utilizing the forum on DropPounds tended to lose weight and then their weight tended to level off. Results exploring if there was a cubic (which could be indicative of weight regain) or quartic (which could be indicative of yo-yo weight loss of regain, loss, and regain) revealed that there was not sufficient variance within the model at later stages to properly test for cubic or quartic effects. Based on this, the unconditional growth model that included a random linear time predictor and a fixed linear and quadratic predictor was used (ie, common metrics among all models in Table 2, excluding word count). In the model whereby only word count was added to the model, results found that word count alone did not predict weight (ie, B=–0.000009, SE 0.000063, P=.13). That said, results indicated a significant model fit improvement when word count was added compared to the unconditional growth model that included quadratic time (ie, unconditional growth model: –2 log-likelihood [–2LL]=621186, Akaike information criterion [AIC]=621198, Bayesian information criterion [BIC]=621240; word count model: –2LL=588920, AIC=588934, BIC=588984). Based on this improved model fit and based on the need to control for overall usage, all LIWC-specific predictor models were compared to the unconditional growth model and the word count model for determination of improved model fit (see Table 2).

Table 2. Mixed models, model fit statistics.

Predictor variables	Unconditional growth	Word count model	Model 1	Model 2	Model 3	Model 4
-2 Log-likelihood	621186	588920	588915	588919	588909	588898
AIC	621198	588934	588935	588937	588933	588919

The word count model represents the unconditional growth model, which includes a random linear effect and a fixed linear and quadratic effect of time, as well as word count per week per participant. Model 1 represents analyses that included topic-relevant words, such as health, ingestion, and motion. Model 2 represent analyses that include positive and negative affect words. Model 3 represents theoretically meaningful words including social process, past-tense verb use, certainty, exclusion, and third-person singular words. Model 4 represents analyses that include conjunction, motion, and past-tense verb use. With regard to –2LL and AIC, lower values indicate better model fit.

Table 3 reports results of the 4 models. In general, results suggest that the content words explored in model 1 (ie, health, ingestion, and biology) and model 2 (ie, positive and negative emotion words) were not predictive of fluctuations in weight from 1 week to the next. Results from the theoretically driven model (model 3) indicated that use of past-tense verbs was predictive of weight loss. Specifically, when individuals utilized more past-tense verbs in their writing (an indication of thinking about the past), they tended to have lower weight that week relative to other weeks. In addition, use of exclusion words was also predictive of fluctuations in weight. Specifically, during weeks when individuals utilized more exclusion words, which in previous work has been linked to increased honesty, there was a statistical trend for those individuals to exhibit higher weight during the same week. The final model, which was more data-driven, found that the 3 LIWC variables with the most variance were all associated with weekly fluctuations in weight. In particular, increased use of past-tense verbs and discussions about motion, which is a content word that could be indicative of physical activity, during the week were associated with lower weights during those weeks. Finally, results indicated that increased use of conjunctions, which has not been linked with any psychological variables in the past, was associated with increased weight during those weeks. Examination of the model fit statistics (ie, –2LL, AIC, and BIC) indicated that model 4 exhibited the greatest overall model fit.

XSL•FO
RenderX

Table 3. Mixed models predicting weekly changes in weight.

Predictor variables	Model 1		Model 2		Model 3		Model 4	
	B (SE)	*P*	B (SE)	*P*	B (SE)	*P*	B (SE)	*P*
Fixed effects								
Fixed intercept	187.1 (0.5011)	<.001	187.1 (0.5011)	<.001	187.1 (0.5011)	<.001	187.1 (0.5011)	<.001
Time	–0.2817 (0.0059)	<.001	–0.2817 (0.0059)	<.001	–0.2817 (0.0059)	<.001	–0.2816 (0.0059)	<.001
Quadratic time	0.0016 (0.0000)	<.001	0.0016 (0.00002)	<.001	0.0016 (0.0000)	<.001	0.0016 (0.0000)	<.001
Word count	–0.0004 (0.0002)	.03	–0.0002 (0.0003)	.62	–0.0001 (0.0004)	.76	–0.0008 (0.0004)	.07
Model 1								
Health	–0.0153 (0.0139)	.27						
Ingestion	–0.0002 (0.0096)	.98						
Biology	0.0112 (0.0106)	.29						
Model 2								
Positive emotion			0.0069 (0.0083)	.41				
Negative emotion			–0.0068 (0.0097)	.48				
Model 3								
Social process					–0.0028 (0.0034)	.40		
Past-tense verbs					–0.0164 (0.0061)	.007		
Certainty					–0.0126 (0.0157)	.42		
Exclusion					0.0148 (0.0082)	.07		
Third-person sing					0.0202 (0.0134)	.13		
Model 4								
Conjunction							0.0203 (0.0063)	.001
Motion							–0.0191 (0.0084)	.02
Past tense							–0.0109 (0.0056)	.05
Random effects								
Linear time	0.121 (0.003)	<.001	0.121 (0.003)	<.001	0.121 (0.003)	<.001	0.121 (0.003)	<.001
Random intercept	1994.6 (31.7)	<.001	1994.5 (31.7)	<.001	1994.5 (31.7)	<.001	1994.5 (31.7)	<.001
Residual	16.4 (0.082)	<.001	16.4 (0.082)	<.001	16.4 (0.082)	<.001	16.4 (0.082)	<.001

Post Text Results

Overview

To give some context to results of the mixed model analyses, some examples of posts that exemplify the different types of utterance usages are provided subsequently.

Past-Tense Verbs

Posts with past-tense verbs use, which was associated with lower weight during weeks when they were used, tended to compare a person's current weight situation to the past. Often, these posts were very long and some even started from childhood and worked forward from there as a sort of "weight loss journey" chronicle. For example, here are excerpts from 1 forum user:

XSL•FO
RenderX

Past Tense Example 1

I have struggled with my weight since third grade. I have tried so many different diets—a childhood friend and I were first introduced to WW when we were in the fifth grade, needless to say it really didn't work...This was the first time in years that I was looking thinner, feeling better about myself and drastically dropping dress sizes—clothes shopping became fun! I got down to 170-175, yeah, I know that still sounds heavy, but I carry my weight differently and looked 20 pounds lighter than I actually was...After graduating, my husband and I thought that our life was starting, we would start our family and everything would just be right. In 2005 we unexpectedly became pregnant...my heaviest weight is 339.6 lb and I am only 5′4″, my goal weight seems so far away at 185 lb. So far, this program has been working well—it is easy to follow, now I just need to exercise.

Past Tense Example 2

I feel you—I think the biggest part of me being apprehensive to this whole weight loss thing again (in terms of food) was counting calories. 1) I was terrified of finding out what I consume, considering I work out pretty meticulously and wasn't losing a thing, and 2) I was daunted at how much work it takes. Eating out becomes a headache, I was scared of having to measure everything, what happens when I just don't KNOW the calories-do I avoid it all together? I fell off the "healthy lifestyle" wagon after few illnesses that happened right after the other. I couldn't exercise and started eating whatever I wanted. But realistically, it does become second nature. I lost 110 the last time I "counted" and only actively counted. Eventually it not only became habit (knowing WHAT to eat and how much) but I started becoming full off of what I NEEDED vs wanted. My dad just knew how many points everything was and knew how to subconsciously eye out portions. A big thing, which I think not too many people think of, is our cravings and their relation to our taste buds and what our body is taking in nutritionally...You can literally train your body into WANTING healthier foods...So, I think, not only do watching calories and what you eat become second nature, but it becomes fulfilling and preferable.

Motion-Related

With regard to the motion-related posts, a scan of the posts with the highest use of motion words tended to revolve around 2 general themes: most commonly a detailed report about an exercise routine, race/run/marathon, and far less often the high motion posts were detailed travel reports on a recent life-changing vacation/trip that inspired weight loss in some way. For example, here are some short excerpts from 1 of the posts reporting on an exercise routine/race:

Motion Exercise Activity Example 1

Yesterday I ran (maybe run is too strong) the St. Jude Memphis Marathon. I've been training for months. If you count the time building up a solid running base, training has actually been almost 2 years. After my first half in April, I felt good afterwards and figured I had it in me to try the full distance. I wanted the feeling of not being sure I could finish the race. The half marathon never made me feel like that. I was told the marathon would deliver. I was told correctly.

Motion Exercise Activity Example 2

What is the most common thing you see women do in the gym? Cardio. And if they do lift weights they pick up a 5-pound dumbbell and do endless reps. As we have discussed, women need to lift heavy, challenging weights just like men in order to gain muscle. While machines do provide sufficient stimulation to gain muscle, nothing can beat free-weight/compound exercises. Now, we will go over a few free-weight and compound exercises that we will incorporate in the weight training presented. [Followed by definitions of each exercise and then a workout schedule]

And here is an example of the beginning of the tale of a "life-changing" trip:

Motion Trip Example 1

So this trip I just took. It was a trip of a lifetime. It ranks up there with my wedding, my brother's wedding in Scotland, and in a way it might be even more significant than even I realize right now. Something happened while on this trip. I turned a corner in this journey. I said goodbye to that girl, the one who has been with me my entire life. I let go of my inner fat girl. I love her, but I don't need her anymore...I felt something inside of me click into place, I felt like I had ARRIVED...

Exclusion-Related

When examining the posts that tended to be related to exclusions (and were associated with slight weight gain), the posts often suggested that the person was providing explanations for a failed plan at weight loss. For example, here are 2 from forum users:

Exclusion Example 1

I lost almost a year to a serious foot injury that was misdiagnosed and then turned out to be an injury that heals slowly. I had been exercising regularly up until then and had lost some weight, but gained it all back plus 15 pounds more during my enforced inactivity. I've just started exercising regularly and have just added [DropPounds] to the equation. I'm not an enthusiastic exerciser either, but I have given myself a few options and I can change it up enough to stick with it. If the gym doesn't work, see what you can do at home. My cable provider has Exercise on Demand; I haven't tried them because that's not really my thing, but I did notice them. If you have cable, you might want to check out what's available. You don't

XSL•FO
RenderX

need a lot of room or fancy equipment. Forget the gym, because if it's that much hassle it's easy to talk yourself out of it. If you don't make it easy for yourself to do, you won't do it.

Exclusion Example 2

Here is an example that happened to me recently. I went home to visit my mom for a long weekend, and ate out a lot. I ate in a reasonable manner, only went overbudget 1 day. I also did a 16-mile run with hard, fast segments in it, and ran at least an hour every day. I was 6 pounds heavier when I got home. My fingers were so swollen I could not get my wedding ring off. After about 4 days of eating food I was accustomed to, I was back to my normal weight. I am convinced it was all the sodium. My mother also salts everything as if she needs to preserve her food before eating it. It's ridiculous...I was once on an 8-week plateau during my weight loss journey. I didn't change anything up; eventually I worked through it. Sometimes just sticking with the plan works. My husband (who has lost 120 pounds) was more of a "cycler" mixing it up if he hit a plateau. He also does not count calories, but focuses on eating whole foods and being aware of his appetite and eating only when he is hungry.

Conjunctions

In a similar vein to exclusion words, use of conjunctions was also associated with modestly higher weight during those weeks. When reading through posts that utilized many conjunctions, responses tended to have a sort of "lessons learned" flavor to them. For example:

Conjunction Example 1

I relearned today NOT [sic] to eat cereal! More on that in a moment. If I run right after I get up and am not starving, I won't eat beforehand. If I'm a little peckish, a handful of cereal will usually do me. If I'm hungry, I'll try a granola bar. If I'm getting up early for a race, I'll usually have scrambled eggs and toast or a granola bar. I also drink water before I go out, even if early in the morning. Did a 3.35-mile run with my boyfriend today, we had to stop halfway through because he felt ill from the coffee and cereal we had for breakfast (I definitely was starting to be a little sick of the milk but was pushing through since it wasn't so bad). I think I might need to just go out on my own for morning runs when I'm visiting him—he likes to eat and then wait until he's digested a bit, and I'd rather just go and come back and be able to take my time getting ready for the day.

Conjunction Example 2

...I feel that pain man, 3 kids and a wife myself! Have made it a priority to get in my exercise, and told the family that as well. I figure a healthier Dad is better than a Fat/Dead from a Heart Attack one, LMAO! My job makes it a bit easier; I can work a couple hours into the middle of the day for exercise. If I had to wait until evening, I'm sure I wouldn't be able to do it consistently, and I am not a morning person.

Discussion

Overall, results of this study suggest that there was a significant impact of time on weight loss for those individuals utilizing the DropPounds forum and tracking system. Beyond that, we also found that the style of utterances within an online weight loss forum is associated with weekly fluctuations in weight. Based on the models we ran, it appears that increased use of past-tense verbs and words related to motion were associated with lower weekly weights whereas increased use of conjunctions (eg, and, but, whereas) and exclusion words (eg, but, without, exclude) were both associated with higher weight during the weeks when these utterances were used more. In general, our results have generally verified the finding that style words have much more predictive power than content words [13].

It is important to be mindful that in the current analyses time was taken into account, but the analyses are largely more akin to a repeated set of cross-sectional analyses as opposed to longitudinal analysis. This is because although weight and posts were measured frequently, the actual associations are being explored when both were measured at the same time. Although it is plausible to create models that allow for temporal relations to be determined (eg, models whereby you lag your predictor such that the previous instance of your predictor predicts the current instance of your outcome variable [35]), this dataset with its frequent missing data is not well suited for this type of lagged analysis. Because this is more akin to cross-sectional analyses, it is impossible to determine the directionality of the effects. With this in mind, the quotation examples appear to provide some insights that seem to imply the likelihood that the posts are likely more driven by the weight loss than the posts driving weight loss. This is based on the fact that all the posts, not surprisingly, have some degree of self-reflection and "sense-making" of weight loss efforts, including behaviors such as exercise.

Even if these results are more indicative of providing explanations for weight loss, they still do provide some insights on the types of activities that individuals utilizing a weight loss forum might want to participate in and be a key target to facilitate within the forums. For example, it appears that there might be some advantage with reflecting on the past to place current efforts at weight loss into perspective based on the association with past-tense verbs. Based on this, a possible future design implication to explore is to facilitate individuals in this sort of reflective "weight loss journey" process. Related to the motion words, it seems logical to conclude that running some type of race and exercise more generally is a valuable strategy to use to foster weight loss. This fits with research suggesting that diet and exercise together appear to be particularly useful for promoting weight loss [36]. From a system design standpoint, a logical strategy to incorporate into online forums could be the facilitation of individuals into partaking races and other exercise routines more actively.

Related to the constructs that were associated with weight gain (ie, exclusion and conjunctions), the previous research on these

XSL•FO
RenderX

variables, particularly exclusion, tends to be associated with increased honesty. It does seem quite plausible that this is simply a reflection of "owning up" to any current weight gain. Because conducting the lagged explorations is not possible in the current dataset based on missing data, it is not possible to examine if these sorts of observations would result in better weight loss the following week. It is plausible, however, that this sort of effect would happen. Regardless, these word types (ie, exclusion and conjunctions) might be valuable identifiers of honesty within the system, which could be used to better understand the current psychological state of users of the forum.

Because this work utilizes secondary data, results suggesting that there was weight loss followed by a leveling off of weight should be considered more descriptive than causative. As already articulated, the analyses controlled for time, but the analyses are more akin to cross-sectional analyses than longitudinal analyses particularly when exploring the relation between utterances and weight. As such, directionality of the effects cannot be determined. Further, the current sample is a convenience sample of real-world users of the DropPounds system. Based on this, there is a natural selection bias in the sample. Despite these limitations, this convenience sample does provide ecological validity in that it is a real-world sample of individuals actively working on weight loss. As such, it is valuable dataset for the specific questions we sought to answer focused on understanding how natural utterances in an online forum are related to fluctuations in weight.

To the best of our knowledge, this study represents the first time styles of utterances within an online forum have been associated with fluctuations in weight. These exploratory results provide empirical evidence for possible interactions to specifically target and facilitate within online weight loss forums, such as participating in races/marathons and reflecting on current weight relative to past experience. In addition, results suggest that certain types of utterances (ie, exclusion in particular) may be a valuable proxy for identifying when a person is making an honest appraisal of their current difficulties with weight. This information could be used to help flag individuals that might need additional help within these online forums. Future work could explore trying to implicate these lessons into the design of more robust online forum interactions.

Acknowledgments

The authors thank Fit Now Inc for providing the data to conduct these analyses.

Conflicts of Interest

None declared.

References

1. Fox S, Duggan M. Health Online. Washington, DC: Pew Internet & American Life Project; 2013 Jan 15. URL: http://www.pewinternet.org/files/old-media//Files/Reports/PIP_HealthOnline.pdf [accessed 2014-07-25] [WebCite Cache ID 6RKtgIaoE]
2. Mo PK, Coulson NS. Exploring the communication of social support within virtual communities: a content analysis of messages posted to an online HIV/AIDS support group. Cyberpsychol Behav 2008 Jun;11(3):371-374. [doi: 10.1089/cpb.2007.0118] [Medline: 18537512]
3. Mo PK, Coulson NS. Empowering processes in online support groups among people living with HIV/AIDS: A comparative analysis of 'lurkers' and 'posters'. Comput Hum Behav 2010 Sep;26(5):1183-1193. [doi: 10.1016/j.chb.2010.03.028]
4. Mo PK, Coulson NS. Developing a model for online support group use, empowering processes and psychosocial outcomes for individuals living with HIV/AIDS. Psychol Health 2012;27(4):445-459. [doi: 10.1080/08870446.2011.592981] [Medline: 21854088]
5. Newman MW, Lauterbach D, Munson SA, Resnick P, Morris ME. It's not that I don't have problems, I'm just not putting them on Facebook: challenges and opportunities in using online social networks for health. In: Proceedings of the ACM 2011 conference on Computer supported cooperative work.: ACM; 2011 Presented at: CSCW'11; March 19-23, 2011; Hangzhou, China p. 341-350. [doi: 10.1145/1958824.1958876]
6. McAlister AL, Perry CL, Parcel GS. How individuals, environments, and health behaviors interact: Social cognitive theory. In: Glanz K, Rimer BK, Viswanath K, editors. Health Behavior and Health Education: Theory, Research, and Practice. San Francisco: Jossey-Bass; 2008.
7. Bennett GG, Glasgow RE. The delivery of public health interventions via the Internet: actualizing their potential. Annu Rev Public Health 2009;30:273-292. [doi: 10.1146/annurev.publhealth.031308.100235] [Medline: 19296777]
8. Eysenbach G, Powell J, Englesakis M, Rizo C, Stern A. Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. BMJ 2004;328(7449):1166.
9. 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]
10. Moyers TB, Martin T, Christopher PJ, Houck JM, Tonigan JS, Amrhein PC. Client language as a mediator of motivational interviewing efficacy: where is the evidence? Alcohol Clin Exp Res 2007 Oct;31(10 Suppl):40s-47s. [doi: 10.1111/j.1530-0277.2007.00492.x] [Medline: 17880345]

XSL•FO
RenderX

11. Moyers TB, Miller WR, Hendrickson SM. How does motivational interviewing work? Therapist interpersonal skill predicts client involvement within motivational interviewing sessions. J Consult Clin Psychol 2005 Aug;73(4):590-598. [doi: 10.1037/0022-006X.73.4.590] [Medline: 16173846]
12. Amrhein PC. How does motivational interviewing work? What client talk reveals. J Cogn Psychother 2004 Oct 01;18(4):323-336. [doi: 10.1891/jcop.18.4.323.64001]
13. Tausczik YR, Pennebaker JW. The psychological meaning of words: LIWC and computerized text analysis methods. Journal of Language and Social Psychology 2009 Dec 08;29(1):24-54. [doi: 10.1177/0261927X09351676]
14. Pennebaker JW. The Secret Life of Pronouns: What Our Words Say About Us. New York: Bloomsbury Press; 2011.
15. Li V, McDonald DW, Eikey EV, Sweeney J, Escajeda J, Dubey G, et al. Losing it online: characterizing participation in an online weight loss community. In: Proc ACM Groupwork (GROUP'14). 2014 Presented at: GROUP'14; November 9-12, 2014; Sanibel Island, FL.
16. Tate DF, Wing RR, Winett RA. Using Internet technology to deliver a behavioral weight loss program. JAMA 2001 Mar 7;285(9):1172-1177. [Medline: 11231746]
17. Womble LG, Wadden TA, McGuckin BG, Sargent SL, Rothman RA, Krauthamer-Ewing ES. A randomized controlled trial of a commercial internet weight loss program. Obes Res 2004 Jun;12(6):1011-1018. [doi: 10.1038/oby.2004.124] [Medline: 15229342]
18. Finley CE, Barlow CE, Greenway FL, Rock CL, Rolls BJ, Blair SN. Retention rates and weight loss in a commercial weight loss program. Int J Obes (Lond) 2007 Feb;31(2):292-298. [doi: 10.1038/sj.ijo.0803395] [Medline: 16755283]
19. Harvey-Berino J, West D, Krukowski R, Prewitt E, VanBiervliet A, Ashikaga T, et al. Internet delivered behavioral obesity treatment. Prev Med 2010 Aug;51(2):123-128 [FREE Full text] [doi: 10.1016/j.ypmed.2010.04.018] [Medline: 20478333]
20. Neve M, Morgan PJ, Collins CE. Weight change in a commercial web-based weight loss program and its association with website use: cohort study. J Med Internet Res 2011;13(4):e83 [FREE Full text] [doi: 10.2196/jmir.1756] [Medline: 21993231]
21. 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]
22. Weinstein PK. A review of weight loss programs delivered via the Internet. J Cardiovasc Nurs 2006;21(4):251-8; quiz 259. [Medline: 16823276]
23. Bensley RJ, Brusk JJ, Rivas J. Key principles in internet-based weight management systems. Am J Health Behav 2010;34(2):206-213. [Medline: 19814600]
24. Pennebaker JW, Chung CK, Ireland M, Gonzales A, Booth RJ. The development and psychometric properties of LIWC. In: LIWC2007 Manual. Austin, TX: LIWC.net; 2007.
25. Pennebaker JW, Booth RJ, Francis ME. Linguistic inquiry and word count: LIWC2001. In: LIWC Operators Manual 2001. Mahwah, NJ: Erlbaum Publishers; 2001:2001.
26. Tausczik YR. Linguistic analysis of workplace computer-mediated communities. In: University of Texas Repository. Austin, TX: University of Texas; 2009.
27. Pennebaker JW, Francis ME. Cognitive, emotional, and language processes in disclosure. Cogn Emot 2010 Sep 10;10(6):601-626. [doi: 10.1080/026999396380079]
28. Rude S, Gortner E, Pennebaker J. Language use of depressed and depression-vulnerable college students. Cogn Emot 2004 Dec;18(8):1121-1133. [doi: 10.1080/02699930441000030]
29. Gunsch MA, Brownlow S, Haynes SE, Mabe Z. Differential forms linguistic content of various of political advertising. JBroadcast Electron Med 2010 Jun 07;44(1):27-42. [doi: 10.1207/s15506878jobem4401_3]
30. Kowalski RM. "I was only kidding!": victims' and perpetrators' perceptions of teasing. Personality and Social Psychology Bulletin 2000 Feb 01;26(2):231-241. [doi: 10.1177/0146167200264009]
31. Simmons R, Gordon P, Chambless D. Pronouns in marital interaction. Psychol Sci 2005 Dec;16(12):932-936. [doi: 10.1111/j.1467-9280.2005.01639.x] [Medline: 16313655]
32. Sexton JB, Helmreich RL. Analyzing cockpit communications: the links between language, performance, error, and workload. Hum Perf Extrem Environ 2000 Oct;5(1):63-68. [Medline: 12190082]
33. Leshed G, Hancock JT, Cosley D, McLeod PL, Gay G. Feedback for guiding reflection on teamwork practices. In: Proc ACM Group Work (Group'07). 2007 Presented at: Group'07; November 4-7, 2007; Sanibel Island, FL p. 217-220. [doi: 10.1145/1316624.1316655]
34. Singer JD, Willett J. Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence. Oxford: Oxford University Press; 2003.
35. Hekler EB, Buman MP, Ahn D, Dunton G, Atienza AA, King AC. Are daily fluctuations in perceived environment associated with walking? Psychol Health 2012;27(9):1009-1020. [doi: 10.1080/08870446.2011.645213] [Medline: 22214492]
36. Curioni CC, Lourenço PM. Long-term weight loss after diet and exercise: a systematic review. Int J Obes 2005 Oct;29(10):1168-1174. [doi: 10.1038/sj.ijo.0803015] [Medline: 15925949]

XSL•FO
RenderX

Abbreviations

AIC: Akaike information criterion
BIC: Bayesian information criterion
LIWC: Language Inquiry and Word Count

Edited by G Eysenbach; submitted 25.07.14; peer-reviewed by A Miller, R Krukowski; comments to author 17.09.14; revised version received 18.09.14; accepted 22.09.14; published 08.12.14.

Please cite as:
Hekler EB, Dubey G, McDonald DW, Poole ES, Li V, Eikey E
Exploring the Relationship Between Changes in Weight and Utterances in an Online Weight Loss Forum: A Content and Correlational Analysis Study
J Med Internet Res 2014;16(12):e254
URL: http://www.jmir.org/2014/12/e254/
doi:10.2196/jmir.3735
PMID:25513997

©Eric B Hekler, Gaurav Dubey, David W McDonald, Erika S Poole, Victor Li, Elizabeth Eikey. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 08.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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

Short Paper

Ethical Issues in Using Twitter for Public Health Surveillance and Research: Developing a Taxonomy of Ethical Concepts From the Research Literature

Mike Conway[1], PhD

University of California San Diego, Department of Family and Preventive Medicine, La Jolla, CA, United States

Corresponding Author:
Mike Conway, PhD
University of California San Diego
Department of Family and Preventive Medicine
9500 Gilman Drive # 0905
La Jolla, CA, 92093
United States
Phone: 1 858 8224478
Fax: 1 858 3001099
Email: mike.conway@utah.edu

Abstract

Background: The rise of social media and microblogging platforms in recent years, in conjunction with the development of techniques for the processing and analysis of "big data", has provided significant opportunities for public health surveillance using user-generated content. However, relatively little attention has been focused on developing ethically appropriate approaches to working with these new data sources.

Objective: Based on a review of the literature, this study seeks to develop a taxonomy of public health surveillance-related ethical concepts that emerge when using Twitter data, with a view to: (1) explicitly identifying a set of potential ethical issues and concerns that may arise when researchers work with Twitter data, and (2) providing a starting point for the formation of a set of best practices for public health surveillance through the development of an empirically derived taxonomy of ethical concepts.

Methods: We searched Medline, Compendex, PsycINFO, and the Philosopher's Index using a set of keywords selected to identify Twitter-related research papers that reference ethical concepts. Our initial set of queries identified 342 references across the four bibliographic databases. We screened titles and abstracts of these references using our inclusion/exclusion criteria, eliminating duplicates and unavailable papers, until 49 references remained. We then read the full text of these 49 articles and discarded 36, resulting in a final inclusion set of 13 articles. Ethical concepts were then identified in each of these 13 articles. Finally, based on a close reading of the text, a taxonomy of ethical concepts was constructed based on ethical concepts discovered in the papers.

Results: From these 13 articles, we iteratively generated a taxonomy of ethical concepts consisting of 10 top level categories: privacy, informed consent, ethical theory, institutional review board (IRB)/regulation, traditional research vs Twitter research, geographical information, researcher lurking, economic value of personal information, medical exceptionalism, and benefit of identifying socially harmful medical conditions.

Conclusions: In summary, based on a review of the literature, we present a provisional taxonomy of public health surveillance-related ethical concepts that emerge when using Twitter data.

(J Med Internet Res 2014;16(12):e290) doi:10.2196/jmir.3617

KEYWORDS

social media; twitter messaging; ethics

Introduction

Since its inception in 2006, the microblog platform Twitter has become a key resource for understanding—and sometimes predicting—mass behavior, particularly in the area of marketing [1] and politics [2]. More recently, the public health community has recognized Twitter's potential for public health surveillance [3,4] with applications including monitoring the prevalence of

XSL•FO
RenderX

infectious diseases in the community [5,6], identifying early-stage disease outbreaks [7], detecting disease outbreaks in mass gatherings [8], and recognizing and understanding health behaviors, like temporal variability in problem drinking [9], and attitudes toward emerging tobacco products such as electronic cigarettes and hookah [10]. Despite the clear utility of using Twitter to augment current public health surveillance, there remains doubt among regulatory authorities, ethics committees, and individual researchers regarding ethically appropriate conduct in this kind of large-scale research, where a single researcher can automatically process hundreds of millions of public tweets. Adding to this difficulty is the fact that many Twitter researchers are based in university computer science and engineering departments, environments that often have not shared as long a tradition of ethical and regulatory oversight as health-related fields.

While there has been significant research effort in developing ethical guidelines for conducting Internet discussion forum-based research generally [11] and for developing ethical guidelines on appropriate use of social media for clinicians [12], there is little current work addressing ethical problems in large-scale automatic Twitter-based public health research. In this paper, we attempt to address this problem by systematically reviewing ethical content in Twitter-based public health surveillance papers with a view to: (1) explicitly identifying a set of potential ethical issues and concerns that may arise when researchers work with Twitter data, and (2) providing a starting point for the formation of a set of best practices for public health surveillance through the development of a taxonomy of ethical concepts derived from the research literature.

Methods

In this review, we are focused on exploring ethical issues that have been identified in published Twitter-based public health surveillance research papers. Relevant research is dispersed across several broad research areas, including biomedicine (Medline), computer science and engineering (Compendex), philosophy (Philosopher's Index), and psychology (PsycINFO). As we are primarily interested in ethical and regulatory issues and how they relate to public health surveillance, with the aid of a biomedical librarian, we designed a complex set of queries for each indexing service to identify those papers that included ethics-related terms in their titles or abstracts, such as "IRB" (institutional review board), "ethics", and "privacy". See Figure 1 for a complete list of keywords.

As the focus of this review is on ethical issues in large-scale automatic Twitter-based research for public health surveillance, we excluded work centered on non-microblog social media platforms (eg, Facebook). We also excluded work on policy and clinician/student professionalism (eg, proposed guidelines for governing clinician interaction with patients via Twitter), and research focused on non-health related topics (eg, marketing) with the exception of those articles concentrating on automatically identifying personality variables from Twitter feeds.

After searching the four databases with queries shown in Figure 1, we screened articles by titles and abstracts, discarding papers that were not available on an open-access basis or via the University of California library system. We began identifying ethical concepts by carefully reading two papers that, through our initial review, we identified as being especially rich in ethical content [13,14]. From these two initial papers, we highlighted sections of the text discussing ethical content and iteratively constructed an initial ethical taxonomy. We then carefully reviewed the remaining 11 papers, adding to and refining the taxonomy. Our methodology was inspired by, but is not identical to, that used by Strech et al [15] who used a rigorous grounded theory methodology to comprehensively investigate ethical issues in the dementia literature. Our aim in this short paper is limited to producing an outline of the ethical issues identified in the Twitter-based public health surveillance research literature.

XSL•FO
RenderX

Figure 1. Literature search queries for PubMed, Compendex, Philosopher's Index, and PsycINFO.

Pubmed Query

```
(((((((((twitter[All Fields] OR tweet[All Fields] OR tweets[All Fields] OR microblog[All
Fields] OR microblogging[All Fields] OR microblogs[All Fields]) AND ("jurisprudence"[mh]
OR "ethics"[mh] OR "comprehension"[mh] OR "human rights"[mh] OR "decision making"[mh] OR
"philosophy"[mh] OR "religion"[mh] OR "morals"[mh] OR "paternalism"[mh] OR "patient
advocacy"[mh] OR "ethics"[All Fields] OR "ethic" [All Fields] OR "ethical" [All Fields]
OR "ethically" [All Fields] OR "beneficence" [All Fields] OR "confidentiality" [All
Fields] OR "confidential" [All Fields] OR "decision Making" [All Fields] OR "informed
consent" [All Fields] OR "mental competence" [All Fields] OR "Autonomy" [All Fields] OR
"justice" [All Fields] OR "moral" [All Fields] OR "morality" [All Fields] OR
"morals" [All Fields] OR "philosophy" [All Fields] OR "religious" [All Fields] OR
"religion" [All Fields] OR "spiritual" [All Fields] OR "prayer" [All Fields] OR
"privacy" [All Fields] OR "private" [All Fields] OR "IRB" [All Fields] OR "institutional
review" [All Fields] OR "right"[All Fields] OR "rights"[All Fields] OR "regulation"[All
Fields] OR "regulations"[All Fields] OR "regulatory"[All Fields] OR "duty"[All Fields] OR
"duties"[All Fields] OR "oversight"[All Fields])))  AND ("2007/01/01"[PDAT] :
"2013/10/25"[PDAT])))))))
```

Retrieved: 66 (1st January 2007 - 25th October2013)
Date of query: 31st October 2013
Service URL: www.pubmed.org

Compendex Query

```
(twitter OR tweet OR tweets OR microblog OR microblogging OR microblogs) AND (twitter OR
tweet OR tweets OR microblog OR microblogging OR microblogs) AND (paternalism OR ethics OR
ethic  OR ethical  OR ethically  OR beneficence  OR confidentiality OR confidential  OR
consent  OR competence  OR Autonomy  OR justice  OR moral  OR morality  OR morals  OR
philosophy OR religious  OR religion  OR spiritual  OR prayer  OR privacy  OR private  OR
IRB  OR institutional  OR duty OR duties OR oversight bioethics OR bioethical OR
bioethical)
```

Retrieved: 219 (1st January 2007 - 25th October 2013)
Date of query: 31st October 2013
Service URL: www.compendex.org

Philosophers Index Query

```
(twitter OR tweet OR tweets OR microblog OR microblogging OR microblogs)
```

Retrieved: 6 (all dates)
Date of query: 31st October 2013
Service URL: www.philindex.org

PsycINFO Query

```
(ALL(twitter) OR ALL(tweet) OR ALL(tweets) OR ALL(microblog) OR ALL(microblogs) OR
ALL(microblogging)) AND  ( ALL(paternalism) OR ALL(ethics) OR ALL(ethic) OR ALL(ethical)
OR ALL(ethically) OR ALL(beneficence) OR ALL(confidentiality) OR ALL(confidential) OR
ALL(consent) OR ALL(competence) OR ALL(Autonomy) OR ALL(justice) OR ALL(moral) OR
ALL(morality) OR ALL(morals) OR ALL(philosophy) OR ALL(religious) OR ALL(religion) OR
ALL(spiritual) OR ALL(prayer) OR ALL(privacy) OR ALL(private) OR ALL(IRB) OR
ALL(institutional) OR ALL(duty) OR ALL(duties) OR ALL(oversight) ALL(bioethics) OR
ALL(bioethical) OR ALL(bioethical))
```

Retrieved: 51 (1st January 2006 - 25th October 2013)
Date of query: 31st October 2013
Service URL: www.psycinfo.org

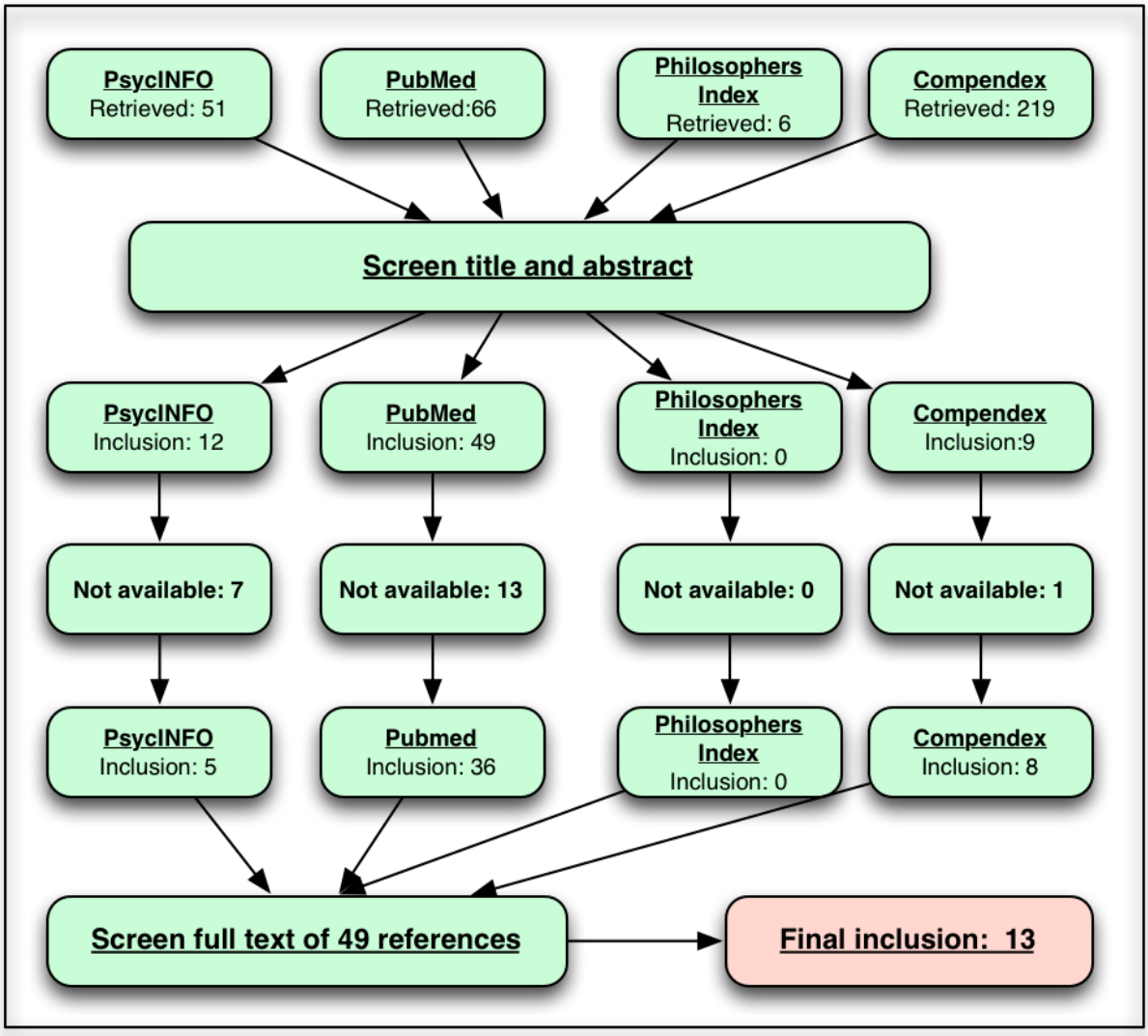
Results

Overview

Our initial set of queries identified 342 references across the four databases (see Figure 2). After title and abstract screening, 49 references remained. After further full-text screening of these 49 references, 13 remained. Five of the papers were from biomedical journals [13,16-19] and six were from computer science and engineering conference proceedings [20-25]. One paper appeared in a journal dedicated to the social and cultural impact of technology [14]. Finally, one paper was published in the proceedings of a collaborative technology conference [26]. All articles were peer reviewed and written in English.

XSL•FO
RenderX

Figure 2. Inclusion/exclusion flowchart.



Taxonomy

Our iterative, taxonomy construction process identified 10 broad ethical categories (eg, privacy, IRB/regulation). Six of these categories included several subcategories (eg, privacy has subcategories including the concept of privacy, IRB/regulation has subcategories including data protection legislation). The most prevalent and complex ethical category we found was privacy with 16 subcategories covering such important ethical concepts as unintended revelation of personal information and population level monitoring vs individual diagnosis (Textbox 1). Multimedia Appendix 1 shows example sentences and paragraphs for each ethical concept in the taxonomy.

XSL•FO
RenderX

Textbox 1. Ethical categories identified during the iterative, taxonomy construction process.

1. Privacy
   - a. Concept of privacy
     - i. Public vs private - doubts about the location of Twitter data on the public/private spectrum [13,14,18,21,24]
     - ii. Fluidity in the concept of privacy - rapidly changing concept of privacy [13,14,18,21,24]
     - iii. Generational differences in the concept of privacy [13,18]
     - iv. Panopticon effect - risk that public health monitoring will change user behavior on Twitter [24]
   - b. Confidentiality
     - i. Data linkage - risk of privacy loss due to linking data from different sources [13,21]
     - ii. Confidentiality - appropriate storage of Twitter data by researchers [14]
     - iii. Right to/desire for anonymity - research using Twitter challenges a participant's right to (and desire for) anonymity [13,14,16]
   - c. Stigmatized medical conditions - concerns about protecting the privacy of those with stigmatized medical conditions (eg, epilepsy, depression) [13,20]
   - d. Twitter's privacy policy - implications of Twitter's privacy policy and how it is understood by users [14,18,19,21,22]
   - e. Twitter is publicly accessible by default - emphasizes that Twitter is a broadcast medium. Unless a user changes privacy settings, tweets are public [14,19,21,22,24]
   - f. Reliability of user provided personal details - reliability of information derived from Twitter when some users use false or whimsical personal details to maintain anonymity [19,21,24]
   - g. Interpreting decontextualized Twitter data as fully representative of users who are in fact multifaceted - possibility that the user might be experimenting with self-presentation or exhibit a belief or behavior in their historical tweets that they no longer adhere to (eg, illegal drug use) [13,24]
   - h. Unintended revelation of personal information - potential for a user to unintentionally provide insights into their mental health, health behaviors, etc, through information garnered from their tweets [20,22-24]
   - i. Personal responsibility of Twitter users - emphasizes the responsibility users have for their posts [22,24]
   - j. Twitter users have no expectation of privacy - emphasizes the researchers' belief that Twitter users have no reasonable expectation of privacy [19,21]
   - k. Identifying users' mental health status or personality traits to:
     - i. identify those in need of treatment [22,25]
     - ii. job placement [20,22-24]
     - iii. targeted marketing [21,23]
     - iv. system interface design (eg, introverts prefer data presented in a certain way) [21,23]
     - v. law enforcement (eg, identifying psychopaths) [22]
   - l. Population level monitoring vs individual diagnosis - difference between using Twitter to identify broad, population level changes and diagnosing individuals [20]
   - m. Potential for discrimination based on health status as garnered from social media [20]
   - n. Danger of inaccurately labeling a user as suffering from a particular health problem [20,22]
   - o. Traceability of Twitter data - risk that tweets can be traced back to the original tweeter if reproduced verbatim in research work, threatening anonymity [13,17]
   - p. Intended audience for tweets - some Twitter users use Twitter as a communication tool for a small group of family and friends and do not expect their tweets to be widely read (ie, hidden in plain sight). Other Twitter users aim to broadcast to the world and gain the maximum number of followers [14,21,24]
2. Informed Consent
   - a. Twitter users are oblivious or unwilling research participants [13]
   - b. Informed consent is difficult (or impossible) to gain (or not required) for large-scale Twitter work [13,14,24]
3. Ethical Theory
   - a. Difficulties in applying current ethical theories to mass Twitter research [13]
   - b. Ethical theories:
     - i. Deontology [13,26]

XSL•FO
RenderX

ii. Utilitarianism [13,26]

iii. Feminism [13]

iv. Communitarianism [13]

v. Application of the "golden rule" [13]

vi. Agile/situational ethics [14]

vii. Rawls' theory of justice [26]

4. IRB/Regulation

a. Citizens' rights to communicate and share information [24,26]

b. Researcher belief that regulatory oversight is not required when using Twitter data [19]

c. Discussion of IRB/ethics committees, generally [14,18]

d. Data protection legislation [14,18]

e. Professional codes of conduct [14]

f. Need for regulatory control, generally [18,20]

g. Privacy regulation by country [14,24]

5. Traditional research vs Twitter research

a. Apomediation - shifting from hierarchical models of research to a situation where the researcher is a potential participant [18]

b. Scale of Twitter-based research - research norms that were developed for small-scale research do not scale to millions of Twitter users [14]

c. Greater distance between researcher and participants - mass Twitter-based research increases the distance between researchers and participants [14]

d. Ambiguous status of participants - the status of participants is more ambiguous than in traditional research (ie, are they consumers, participants, patients, service users, journalists, etc) [14]

e. Increase in researcher power - in mass Twitter research, a single researcher has access to millions of Twitter users, hence increasing researcher power [14]

6. Geographical Information

a. Tracking physical location - potential loss of privacy in tracking Twitter users' physical locations [13,14,19,21]

b. Appropriate geographical granularity – potential loss of privacy in reporting a Twitter user's precise location, compared to their general location. For example, reporting that a Twitter user is *somewhere* in Los Angeles is very different to reporting their precise location in Los Angeles [14]

7. Researcher Lurking [13,14]

8. Economic Value of Personal Information [14]

9. Medical Exceptionalism - health-related matters are qualitatively different from other, non-medical areas and require special attention (and perhaps regulation) [25]

10. Benefit of Identifying Socially Harmful Medical Conditions [22]

Normative Rules

In several of the papers under review, explicit normative rules were presented (or suggested) for conducting and reporting mass Twitter public health surveillance work (Textbox 2). Note that these rules are *discussed* but not necessarily *endorsed*.

XSL•FO
RenderX

Textbox 2. Explicit normative rules for mass Twitter public health surveillance work.

1. When reporting research, avoid quoting directly from users' Twitter streams. Paraphrases should be used [13].
2. Informed consent should be gained from participants [13].
3. Metadata (usernames, location data, etc) should not be disclosed [13].
4. Twitter-based work is human subjects research and requires that some form of appropriate IRB/ethical review take place [14].
5. Data collection should be logged and justified [14].
6. There should be parity between the researcher and participants (ie, the researcher's tweets and their associated locations, if appropriate, should be public) [14].
7. Employment-related profiling for mental health conditions should only be performed in exceptional circumstances (eg, security critical roles) [24].
8. Consent should be gained from potential employees before employment-related profiling for mental health conditions is performed [24].

Discussion

The main output of this research is a taxonomy of ethical concepts derived from close reading of the literature. The taxonomy will be used to help frame future interview-based qualitative research focused on Twitter users' attitudes to the use of microblog data for public health surveillance and, in due course, inform the generation of a set of ethical guidelines for using Twitter for public health surveillance and research. We found that ethical theory was rarely mentioned in the reviewed papers and, when it was discussed, that discussion was typically brief. Only two papers [13,26] explicitly discuss the application of traditional ethical theories (eg, deontology, utilitarianism) to mass Twitter-based public health surveillance. As expected, the bulk of the ethical concepts we discovered were concerned with privacy [13,14,16-25], including frequent references to the fluid and changing nature of the concept of privacy [13,14,18,21,24], and more concretely, to Twitter's privacy policies [14,18,19,21,22]. Discussion of IRBs and regulation (or the lack thereof) was also widespread in the literature [14,18-20,24,26]. Some topics were raised by a single research paper, for example, the idea that the ability to automatically process millions of tweets increases researcher power compared to traditional research methodologies [14], and the idea that the benefits of using Twitter for public health purposes are so great that they mitigate any ethical doubts that would apply to other, non-health-related uses of Twitter data (eg, for commercial gain) [25].

Although inspired by the ethics-oriented qualitative literature review methodology proposed by Strech [15], the approach taken in this review is substantially different, in particular the use of a single reviewer (author MC) rather than a group of reviewers, and the use of *close reading* in place of a theoretically grounded qualitative methodology. A further characteristic of this review is that our search strategy was confined to papers indexed in PubMed, Compendex, PsycINFO, and the Philosopher's Index. Papers that were not available via the University of California San Diego library system or on an open-access basis were excluded. It is likely that we "missed" relevant papers in business disciplines or in those computer science and engineering conferences and journals not indexed by Compendex. However, our purpose in this review was the identification of a broad taxonomy of ethical concepts relevant to Twitter-based public health research using a systematic, reproducible methodology and thus comprehensiveness, while desirable, is not a necessity.

In conclusion, this short paper provides a taxonomy of ethical concepts derived from the research literature. Future work will involve interview-based qualitative research exploring Twitter users' attitudes toward the mining of their data for public health purposes, and ultimately the formation of best practice guidelines for public health surveillance using Twitter data.

Acknowledgments

We would like to thank Dr Dan O'Connor (Head of Medical Humanities at the Wellcome Trust) and Dr Samantha Hurst in the Department of Family and Preventive Medicine at the University of California San Diego for advice and guidance in the design and execution of this research. This work was supported by a grant from the National Library of Medicine K99LM011393.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Taxonomy.

[PDF File (Adobe PDF File), 99KB - jmir_v16i12e290_app1.pdf]

XSL•FO
RenderX

References

1. Bulearca M, Bulearca S. Twitter: a viable marketing tool for SMEs. Global Business and Management Research: An International Journal 2012;2(4):296-309 [FREE Full text]
2. Tumasjan A, Sprenger T, Sandner P, Welpe I. Predicting elections with Twitter: what 140 characters reveal about political sentiment. In: Proceedings of the Fourth International Conference on Weblogs and Social Media. 2010 Presented at: Fourth International Conference on Weblogs and Social Media; 2010; Washington, DC p. 178-185.
3. 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]
4. Stoové MA, Pedrana AE. Making the most of a brave new world: opportunities and considerations for using Twitter as a public health monitoring tool. Prev Med 2014 Jun;63:109-111. [doi: 10.1016/j.ypmed.2014.03.008] [Medline: 24632229]
5. Collier N, Son NT, Nguyen NM. OMG U got flu? Analysis of shared health messages for bio-surveillance. J Biomed Semantics 2011;2 Suppl 5:S9 [FREE Full text] [doi: 10.1186/2041-1480-2-S5-S9] [Medline: 22166368]
6. 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]
7. Collier N, Doan S. Syndromic classification of Twitter messages. In: Szomszor M, Kostkova P, editors. Electronic Healthcare. Berlin: Springer; 2012:186-195.
8. Yom-Tov E, Borsa D, Cox IJ, McKendry RA. Detecting disease outbreaks in mass gatherings using Internet data. J Med Internet Res 2014;16(6):e154 [FREE Full text] [doi: 10.2196/jmir.3156] [Medline: 24943128]
9. West JH, Hall P, Hanson C, Prier K, Giraud-Carrier C, Neeley E, et al. Temporal variability of problem drinking on Twitter. OJPM 2012;02(01):43-48. [doi: 10.4236/ojpm.2012.21007]
10. Myslín M, Zhu SH, Chapman W, Conway M. Using Twitter to examine smoking behavior and perceptions of emerging tobacco products. J Med Internet Res 2013;15(8):e174 [FREE Full text] [doi: 10.2196/jmir.2534] [Medline: 23989137]
11. Eysenbach G, Till JE. Ethical issues in qualitative research on internet communities. BMJ 2001 Nov 10;323(7321):1103-1105 [FREE Full text] [Medline: 11701577]
12. 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;15(8):e184 [FREE Full text] [doi: 10.2196/jmir.2708] [Medline: 23985172]
13. McKee R. Ethical issues in using social media for health and health care research. Health Policy 2013 May;110(2-3):298-301. [doi: 10.1016/j.healthpol.2013.02.006] [Medline: 23477806]
14. Neuhaus F, Webmoor T. Agile ethics for massified research and visualization. Information, Communication & Society 2012 Feb;15(1):43-65. [doi: 10.1080/1369118X.2011.616519]
15. Strech D, Mertz M, Knüppel H, Neitzke G, Schmidhuber M. The full spectrum of ethical issues in dementia care: systematic qualitative review. Br J Psychiatry 2013 Jun;202:400-406 [FREE Full text] [doi: 10.1192/bjp.bp.112.116335] [Medline: 23732935]
16. Sugawara Y, Narimatsu H, Hozawa A, Shao L, Otani K, Fukao A. Cancer patients on Twitter: a novel patient community on social media. BMC Res Notes 2012;5:699 [FREE Full text] [doi: 10.1186/1756-0500-5-699] [Medline: 23270426]
17. Heaivilin N, Gerbert B, Page JE, Gibbs JL. Public health surveillance of dental pain via Twitter. J Dent Res 2011 Sep;90(9):1047-1051 [FREE Full text] [doi: 10.1177/0022034511415273] [Medline: 21768306]
18. O'Connor D. The apomediated world: regulating research when social media has changed research. J Law Med Ethics 2013;41(2):470-483. [doi: 10.1111/jlme.12056] [Medline: 23802898]
19. Burton SH, Tanner KW, Giraud-Carrier CG, West JH, Barnes MD. "Right time, right place" health communication on Twitter: value and accuracy of location information. J Med Internet Res 2012;14(6):e156 [FREE Full text] [doi: 10.2196/jmir.2121] [Medline: 23154246]
20. Sumner C, Byers A, Boochever R, Park GJ. Predicting dark triad personality traits from Twitter usage and a linguistic analysis of tweets. In: Proceedings of the 11th IEEE International Conference on Machine Learning and Applications. 2012 Presented at: 11th IEEE International Conference on Machine Learning and Applications; 2012; Florida. [doi: 10.1109/ICMLA.2012.218]
21. Quercia D, Kosinski M, Stillwell D, Crowcroft J. Our Twitter profiles, our selves: Predicting personality with Twitter. In: Proceedings of the IEEE International Conference on Privacy, Security, Risk and Trust. 2011 Presented at: IEEE International Conference on Privacy, Security, Risk and Trust (PASSAT); 2011; Massachusetts p. 180-185.
22. Wald R, Khoshgoftaar TM, Napolitano A, Sumner C. Using Twitter content to predict psychopathy. In: Proceedings of the 11th IEEE International Conference on Machine Learning and Applications. 2012 Presented at: 11th IEEE International Conference on Machine Learning and Applications; 2012; Florida p. 394-401. [doi: 10.1109/ICMLA.2012.228]
23. Golbeck J, Robles C, Edmondson M, Turner K. Predicting personality from Twitter. In: Proceedings of the IEEE International Conference on Privacy, Security, Risk and Trust (PASSAT). 2011 Presented at: IEEE International Conference on Privacy, Security, Risk and Trust (PASSAT); 2011; Massachusetts p. 149-156. [doi: 10.1109/PASSAT/SocialCom.2011.33]

XSL•FO
RenderX

24. Kandias M, Galbogini K, Mitrou L, Gritzalis D. Insiders trapped in the mirror reveal themselves in social media. In: Network and System Security: Proceedings of the 7th International Conference on Network and System Security. Berlin: Springer; 2013 Presented at: 7th International Conference on Network and System Security; 2013; Madrid p. 220-235. [doi: 10.1007/978-3-642-38631-2]
25. De Choudhury M, Counts S, Horvitz E. Predicting postpartum changes in emotion and behavior via social media. In: Proceedings of the 31st Annual CHI Conference on Human Factors in Computing Systems. 2013 Presented at: 31st Annual CHI Conference on Human Factors in Computing Systems; 2012; France p. 3267-3276. [doi: 10.1145/2470654.2466447]
26. Heverin T. Ethical concerns of Twitter use for collective crisis response. In: Proceedings of the 12th International Conference on Collaboration Technologies and Systems. 2011 Presented at: 12th International Conference on Collaboration Technologies and Systems; 2011; Colorado p. 625-626.

Abbreviations

IRB: institutional review board

Edited by G Eysenbach; submitted 18.06.14; peer-reviewed by E Vayena, HP Lee; comments to author 19.08.14; accepted 28.10.14; published 22.12.14.

Please cite as:
Conway M
Ethical Issues in Using Twitter for Public Health Surveillance and Research: Developing a Taxonomy of Ethical Concepts From the Research Literature
J Med Internet Res 2014;16(12):e290
URL: http://www.jmir.org/2014/12/e290/
doi:10.2196/jmir.3617
PMID:25533619

©Mike Conway. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 22.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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

Impact of an Online Medical Internet Site on Knowledge and Practice of Health Care Providers: A Mixed Methods Study of the Spinal Cord Injury Rehabilitation Evidence Project

Janice J Eng[1,2,3], PT, PhD; Vanessa K Noonan[4], PhD; Andrea F Townson[2,3,5], MD, FRCPC; Caroline E Higgins[6,7], MISt; Jess Rogers[6,7], BA; Dalton L Wolfe[8,9], PhD

[1]Department of Physical Therapy, University of British Columbia, Vancouver, BC, Canada
[2]International Collaboration on Repair Discoveries (ICORD), Vancouver, BC, Canada
[3]Rehabilitation Research Program, GF Strong Rehab Centre, Vancouver, BC, Canada
[4]Rick Hansen Institute (RHI), Vancouver, BC, Canada
[5]Division of Physical Medicine & Rehabilitation, University of British Columbia, Vancouver, BC, Canada
[6]Centre for Effective Practice, Toronto, ON, Canada
[7]Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada
[8]Lawson Health Research Institute, London, ON, Canada
[9]Western University, London, ON, Canada

Corresponding Author:
Janice J Eng, PT, PhD
Department of Physical Therapy
University of British Columbia
Friedman Building
212-2177 Wesbrook Mall
Vancouver, BC, V6T 1Z3
Canada
Phone: 1 604 714 4105
Fax: 1 604 714 4168
Email: janice.eng@ubc.ca

Abstract

Background: It is not known whether ongoing access to a broad-based Internet knowledge resource can influence the practice of health care providers. We undertook a study to evaluate the impact of a Web-based knowledge resource on increasing access to evidence and facilitating best practice of health care providers.

Objective: The objective of this study was to evaluate (1) the impact of the Spinal Cord Injury Rehabilitation Evidence (SCIRE) project on access to information for health care providers and researchers and (2) how SCIRE influenced health care providers' management of clients.

Methods: A 4-part mixed methods evaluation was undertaken: (1) monitoring website traffic and utilization using Google Analytics, (2) online survey of users who accessed the SCIRE website, (3) online survey of targeted end-users, that is, rehabilitation health care providers known to work with spinal cord injury (SCI) clients, as well as researchers, and (4) focus groups with health care providers who had previously accessed SCIRE.

Results: The online format allowed the content for a relatively specialized field to have far reach (eg, 26 countries and over 6500 users per month). The website survey and targeted end-user survey confirmed that health care providers, as well as researchers perceived that the website increased their access to SCI evidence. Access to SCIRE not only improved knowledge of SCI evidence but helped inform changes to the health providers' clinical practice and improved their confidence in treating SCI clients. The SCIRE information directly influenced the health providers' clinical decision making, in terms of choice of intervention, equipment needs, or assessment tool.

Conclusions: A Web-based knowledge resource may be a relatively inexpensive method to increase access to evidence-based information, increase knowledge of the evidence, inform changes to the health providers' practice, and influence their clinical decision making.

XSL•FO
RenderX

(J Med Internet Res 2014;16(12):e296) doi:10.2196/jmir.3453

KEYWORDS

Internet; e-health; spinal cord injury; rehabilitation; evidence-based practice

Introduction

eHealth has been defined as "health services and information delivered or enhanced through the Internet and related technologies" [1]. Online health information was one of the earliest (and remains one of the most frequent) applications of eHealth [2,3]. A population-based European study (over 14,000 participants surveyed) estimated that over 50% of adults use the Internet as a source of consumer health information [4]. More recently, health care professionals have increased their use of the Internet to access medical and health information to update their practice and ultimately improve the outcomes of their patients [3]. While free access to academic electronic databases (eg, PubMed) has increased over the years, health care providers have reported numerous barriers to their use, including inability to access the full text of the articles, lack of skills in searching and appraising the literature, and lack of time to compile all of the relevant evidence [5-7]. The amount of information to answer any one clinical question can be massive, and the literature is often conflicting with various levels of quality. For example, a search on the narrow topic of the effectiveness of anticonvulsant medication for the management of pain following spinal cord injury (SCI) results in 13 studies ranging from randomized controlled trials to pre-post studies and retrospective chart reviews [8].

Formal assessments of the body of scientific evidence are an important and time-saving resource for clinicians wishing to incorporate evidence into their clinical decision making and are a key element of putting knowledge into action (ie, knowledge translation) [9]. The Knowledge-to-Action Cycle is a conceptual framework that builds on more than 30 planned action theories and comprises a dynamic process between a knowledge creation and a knowledge application domain [10]. Three levels or "generations" of information have been identified within the knowledge creation domain [11]. Graham et al [11] refer to the information associated with numerous primary studies of varied quality as the "first generation knowledge" or "knowledge inquiry" stage. Synthesized literature may overcome some of the barriers to compiling and weighting relevant literature [12]. This synthesized or aggregated literature often takes the form of a systematic review or meta-analysis and is referred to as "second-generation knowledge" or "knowledge synthesis" [11]. "Third-generation knowledge" or "knowledge tools" take the knowledge synthesis one step further and present user-friendly information with the intent of influencing practice [11]. These include tools and knowledge products such as practice guidelines, care pathways, or websites that synthesize and provide expert commentary on the literature with regard to current practice [11]. Knowledge tools are often preferred by clinicians as they provide explicit, evidence-based descriptions of the benefits and risks of an intervention at the level of the user [13].

Health providers have long cited that having the evidence in one place would facilitate the translation of research to practice [14]. Several websites targeted at frontline health care providers serve as a knowledge product to compile, synthesize, and update selected literature and to provide expert commentary over a focused topic area (eg, specific disease). For example, the Rehabilitation Measures website assists clinicians in their outcome measure selection by reviewing the psychometric properties of outcome measures and providing instructions for their administration in the rehabilitation field [15], while the disease-specific Stroke Engine website critically appraises stroke-related treatments and outcome measures [16]. Profession-specific sites such as the American College of Physicians Journal Club are designed to filter the literature and deliver expert summaries of articles that warrant immediate attention by physicians in internal medicine [17]. Initially disseminated through a traditional journal format, the content is now complemented with an online version.

Access to timely quality evidence is the first step to rectifying the consistent failure to translate research into practice [18] and reducing the unacceptably long time it takes to translate this evidence [19]. While it is known that delivery of specific Internet-based learning activities that require formal registration and participation (eg, virtual anatomy course, teleconference on a specific client case) can have positive effects on educational outcomes for health care providers and students [20], it is not known whether ongoing access to a broad-based Internet knowledge resource can influence the practice of health care providers. We undertook a mixed methods evaluation of the Spinal Cord Injury Rehabilitation Evidence (SCIRE) Project website to answer this question.

The SCIRE Project is an online resource that provides a synthesis of the evidence underlying SCI rehabilitation interventions and is designed to enable relevant decision making in practice settings [21] and public policy. The target audience are health care professionals, researchers, and policy makers. SCIRE uses transparent methods to synthesize all levels of evidence (from randomized controlled trials to case reports), as well as present an overview of relevant outcome measures (eg, purpose of measure, psychometric properties). Over 70 clinicians and researchers with a high degree of expertise in the topics provide summaries of the evidence with implications for practice and decision making.

The SCIRE Project consists of two main sections. The first is an overview of the rehabilitation evidence for 27 different sections. The specific sections are illustrated in Figure 1. Each section includes background information on the topic, tables of relevant individual studies, a discussion synthesizing the data, a summary of the levels of evidence, and a summary of key points. Standardized search methods and assignment of the levels of evidence are undertaken to ensure that the methods are replicable and transparent. The second section is an overview of 105 different outcome measures related to spinal cord injury.

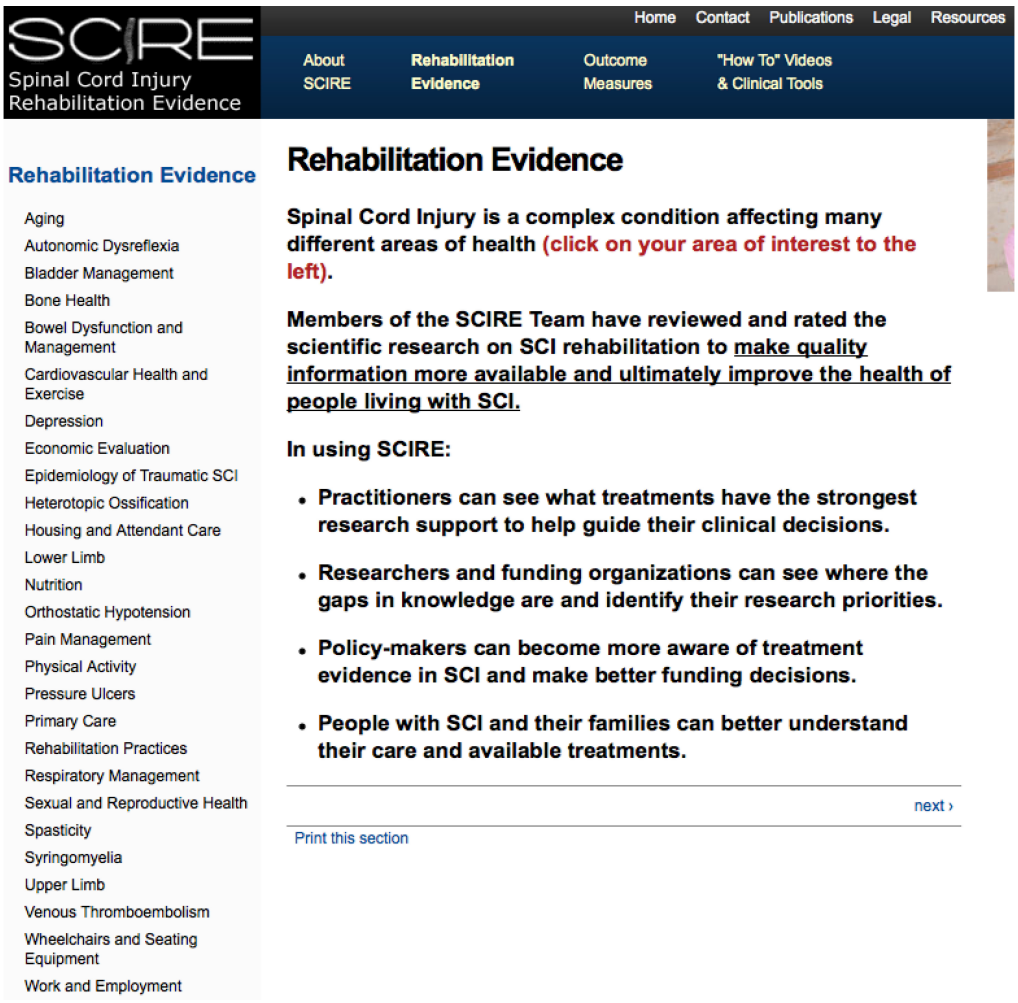
XSL•FO
RenderX

For each outcome measure, a clinical summary including the psychometrics (reliability, validity) and how-to-use section are provided. One outcome measure is illustrated in Figure 2.

The SCIRE Project is relevant to the knowledge creation domain of the knowledge-to-action cycle and creates "third-generation knowledge" or "knowledge tools" to convey user-friendly information with the intent of influencing practice. Numerous gaps have been identified between current practice and best practice in SCI rehabilitation including the underutilization of standardized outcome measures [22], as well as the lack of implementation of the best evidence for SCI-related complications such as urinary tract infections [23] and pressure ulcers [24]. The SCIRE Project has been online since December 2009.

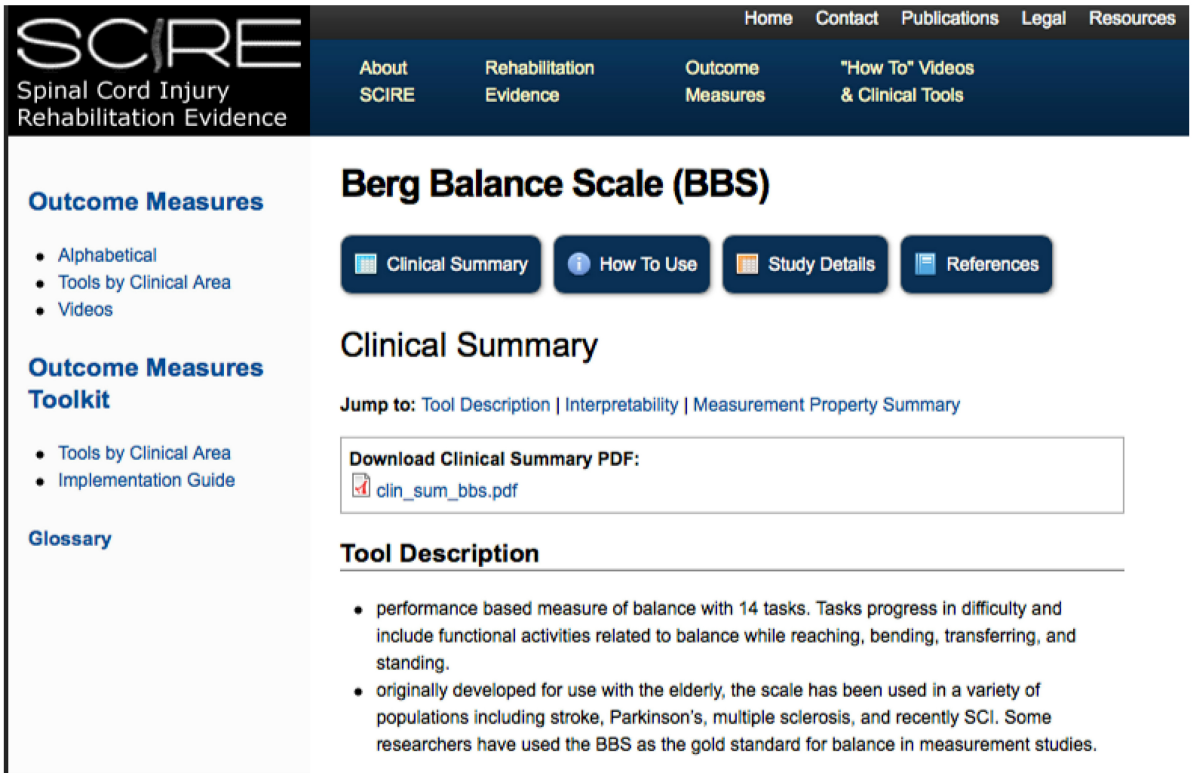
Our objective was to evaluate the impact of the SCIRE Project on the uptake of knowledge. We hypothesized that this Web-based source of information would (1) increase access to evidence-based information for health care providers and researchers and (2) influence the self-report practice of these health care providers in the management of their clients.

Figure 1. Screenshot illustrating the 27 sections providing rehabilitation evidence related to spinal cord injury. Each section includes background information on the topic, tables of relevant individual studies, a discussion synthesizing the data, a summary of the levels of evidence, and a summary of key points.



XSL•FO
RenderX

Figure 2. Screenshot illustrating one of the available 105 outcome measures related to spinal cord injury. For each outcome measure, a clinical summary and how-to-use section is provided.



Methods

Overview

A 4-part mixed methods evaluation was used to evaluate the SCIRE Project. The protocol was approved by the local university ethics board, and participant consent was required for the surveys and focus groups. An interdisciplinary project advisory team of 12 researchers and clinicians in the field of SCI rehabilitation, including the professions of medicine, health care administration and evaluation, physical therapy, occupational therapy, and research provided input to the development of the tools (surveys and interview guides). To assess the SCIRE Project as a knowledge translation tool, questions were developed using Pathman's Awareness-to-Adherence Model [25]. Survey instruments were developed using Dillman's Tailored Design Method [26]. Two re-iterations of each tool were reviewed by the project advisory team beyond the initial version to test logic and functionality. Our study is a first step in determining whether ongoing access to a broad-based Internet knowledge resource can influence the practice of health care providers and utilizes self-report measures that are useful to reveal behaviors and perceptions of clinicians. However, we acknowledge that such measures are subject to various forms of bias (eg, recall bias, social desirability bias).

Part 1: Website Traffic

We monitored website traffic and utilization using Google Analytics from the initial hosting of the website (December 2009) to July 2012. The purpose was to analyze the website visits and visitor behavior to provide a picture of activity on the website.

Part 2: Website Survey

A brief (5 minutes) online survey was hosted on the SCIRE Project website over a 5-week period (July 20-Aug 24, 2012) and appeared when visitors exited the website. The purpose was to supplement the Google Analytics data with an 11-question online survey of which 4 questions related to demographics, 1 question on how they heard about SCIRE, 5 questions on use and usability (including navigation; not reported in this paper), and 1 question on impact (the focus of this paper). The question on impact queried the impact of the SCIRE Project on the participant's access to SCI information and on their clinical practice rated on a 5-point Likert scale.

Part 3: Targeted End-User Survey

The purpose of the end-user survey was to provide an indication of the awareness of the SCIRE Project, and the impact on clinical or research practice. In contrast to the website survey, participants were health care providers or researchers confirmed to work in the field of SCI rehabilitation. The SCIRE Project was developed with the health provider as the main end-user; however, it is recognized that researchers also constitute a major end-user. A sample of convenience was generated by sending the survey link to 22 coordinators associated with SCI rehabilitation centers in Canada who disseminated it to frontline clinicians who worked in SCI rehabilitation (July 26-Sept 7, 2012). An additional 95 surveys were sent to names gathered from Canadian public health care registries when they specified the interest of the provider to be SCI.

The targeted end-user survey queried the demographics of the individual, reasons why they visited the SCIRE Project, how useful was the SCIRE Project to their information needs, and

XSL•FO
RenderX

questions related to the access to information and impact on practice rated on a 5-point Likert scale. Descriptive analyses were used to tabulate the frequency counts to summarize the data for the targeted end-user survey, as well as the website survey.

Part 4: Focus Groups

Online focus groups with health care providers were undertaken to explore how the SCIRE Project website information is applied in the clinical setting. The interview guide included questions on experience with the SCIRE Project website, motivations for using the SCIRE Project, and extent to which the SCIRE Project has impacted practice. Inclusion criteria for participation in the focus groups were health care providers who had used the SCIRE Project previously. They were selected from those in Part 3 who had agreed to be contacted for a focus group. Two online focus groups (n=8, n=7) were conducted and were led by a trained facilitator with technical support. Each focus group session was recorded and transcribed verbatim. Qualitative content analysis was then applied to the transcripts [27]. The transcripts were read and reviewed separately by 2 researchers. The text was coded into meaning units (statements that relate to the same central meaning), and then the codes were grouped into categories (grouping of codes) until no new categories emerged. Categories were then aggregated into themes (concepts that cut across codes and categories and contributed to our understanding of the impact of the SCIRE Project), and relationships among these were determined. Representative quotes were used to illustrate the themes and to assist the readers in judging the transferability of the findings.

Results

Overview

A total of 332 people participated in the evaluation across the website survey, end-user survey, and focus groups.

Part 1: Website Traffic

A total of 195,519 unique visitors accessed the website between December 2009 and July 2012. The majority of website visitors were from (in descending order) United States, Canada, United Kingdom, Australia, and India. The majority of website visitors (77.62%) were new visitors, while 22.38% were returning visitors. Of the top five most visited content pages, four were specific outcome measures (Functional Independence Measure, Symptom Checklist, Jebsen Hand Test, Hospital Anxiety and Depression Scale), while one was a page within the Respiratory Interventions Chapter (Intermittent Positive Pressure Breathing). Content pages involving the outcome measures were viewed 58.56% of the time, while content from the intervention chapters were viewed the other 41.44% of the time.

Part 2: Website Survey

There were 171 individuals who completed the website survey. While 26 countries participated, the majority of responses were from the United States (29.8%, 51/171), Canada (29.8%, 51/171), and Australia (14.0%, 24/171). The respondents were primarily health care providers (59.1%, 101/171), while researchers (26.37%, 45/171), SCI consumers and their families (8.2%, 14/171), policy makers (1.7%, 3/171), and others (4.7%, 8/171) made up the remaining groups. Of the health care providers, the majority were physical therapists (41.6%, 42/101), physicians (23.8%, 24/101), occupational therapists (11.9%, 12/101), and nurses (10.9%, 11/101). The health care providers worked in rehabilitation settings (72.3%, 73/101), community settings (13.9%, 14/101), acute settings (7.9%, 8/101), and other settings (5.9%, 6/101) and had a mean of 6.5 (SD 7.6) years of clinical experience. Researchers had a mean of 5.3 (SD 7.2) years of research experience. Most of the health care provider respondents first heard about SCIRE from an Internet search engine result (46.5%, 47/101), referral from a colleague (31.7%, 32/101), published article (10.9%, 11/101), or conference/workshop (10.9%, 11/101).

The responses of the health care providers (n=101) suggested that the website may have had positive impact on practice (Table 1): 89% agreed or strongly agreed that the SCIRE Project had increased their access to SCI evidence, 91% agreed or strongly agreed that the SCIRE Project had improved their knowledge of SCI evidence, 81% agreed or strongly agreed that the SCIRE Project had helped inform changes to their clinical practice, and 69% agreed or strongly agreed that the SCIRE Project had improved their confidence in treating SCI clients.

The researchers agreed or strongly agreed (82-87%) that the SCIRE Project had increased their access to SCI evidence, improved their knowledge of SCI evidence, and supported their current research activities.

XSL•FO
RenderX

Table 1. Website survey results showing impact of SCIRE on health care providers and researchers (101 health care providers and 45 researchers completed the survey, but some marked an item as not applicable and were not included in the percent calculation, eg, a clinician who had no research responsibilities for item 2).

	n	Strongly agree, %	Agree, %	Undecided, %	Disagree, %	Strongly disagree, %
Health care providers						
SCIRE has improved my knowledge of SCI evidence	98	56.1	34.7	7.1	1.0	1.0
SCIRE supports my current research activities	86	41.8	41.8	14.0	1.2	1.2
SCIRE has helped inform changes to my clinical practice	92	29.3	52.2	15.2	2.2	1.1
SCIRE has increased my confidence in treating SCI clients	90	29.9	38.9	24.4	5.6	1.1
SCIRE has increased my access to SCI evidence	97	58.8	29.9	10.3	0	1.0
SCIRE has not impacted my clinical practice or research activities	98	3.1	5.1	12.2	48.0	31.6
Researchers						
SCIRE has improved my knowledge of SCI evidence	43	44.3	42.0	13.7	0	0
SCIRE supports my current research activities	39	43.6	38.5	15.4	2.5	0
SCIRE has increased my access to SCI evidence	39	59.0	28.2	12.8	0	0
SCIRE has not impacted my research activities	35	0	0	20.0	45.7	34.3

Part 3: Targeted End-User Survey

Of the 146 targeted end-users, the majority of responses resulted from the distribution to the 22 rehabilitation centers (n=111) and the remaining (n=35) from the direct email request. The majority were working in Canada (83.6%, 122/146), with a small proportion (12.3%, 18/146) from the United States and from other countries (4.1%, 6/146). Due to privacy issues, we were not able to directly email the staff at the rehabilitation centers, and thus do not have information about the response rate for this group. Of the 95 direct email requests, 64 opened the email, and 35 completed the survey.

Of the targeted end-users, 69.9% (102/146) were aware of the SCIRE Project. For the users that were aware of the SCIRE Project, 52.0% (53/102) worked in a rehabilitation setting, 30.4% (31/102) in an acute setting, 15.7% (16/102) in a community setting, and 2.0% (2/102) indicated “other”. The respondents were primarily health care providers (66.7%, 68/102), while researchers (22.5%, 23/102) made up the second largest proportion. Of the health care providers, there was representation from physical therapists (32%, 22/68), physicians (16%, 11/68), occupational therapists (10%, 7/68), nurses (24%, 16/68), psychologists (4%, 3/68), and social worker/counselors (4%, 3/68). Health care providers had a mean 8.5 (SD 7.1) years of experience while researchers had 11.5 (10.8) years of experience. The users that were not aware of the SCIRE Project (30.1%, 44/146) had a similar mix of health providers; however, a larger proportion (52%, 23/44) worked in an acute setting.

Of the users who were aware of the SCIRE Project, 62 health care providers and 31 researchers had accessed the website previously, and their feedback is shown in Table 2. Similar to the website survey, the health care providers indicated positive impact of the website on practice (Table 2): 88% agreed or strongly agree that the SCIRE Project had improved their knowledge of SCI evidence, 93% agreed or strongly agreed that the SCIRE Project had increased their access to SCI evidence, 66% agreed or strongly agreed that the SCIRE Project helped inform changes to practice, and 62% agreed or strongly agreed that the SCIRE Project increased their confidence in treating SCI clients.

Of the researchers, a majority (70-81%) agreed or strongly agreed that the SCIRE Project had increased their access to SCI evidence, improved their knowledge of SCI evidence, and supported their current research activities.

XSL•FO
RenderX

Table 2. Targeted user survey results showing impact of SCIRE on health care providers and researchers (62 health care providers and 31 researchers were surveyed, but some marked the item as not applicable and were not included in the percent calculation, eg, a clinician who had no research responsibilities for item 2).

	n	Strongly agree, %	Agree, %	Undecided, %	Disagree, %	Strongly disagree, %
Health care providers						
SCIRE has improved my knowledge of SCI evidence	59	47.1	41.5	9.4	1.9	0
SCIRE supports my current research activities	51	13.0	58.7	26.1	0	2.2
SCIRE has helped inform changes to my clinical practice	56	31.4	37.2	25.5	5.9	0
SCIRE has increased my confidence in treating SCI clients	58	30.8	36.5	26.9	5.8	0
SCIRE has increased my access to SCI evidence	61	50.9	41.9	7.2	0	0
SCIRE has not impacted my clinical practice or research activities	57	21.1	15.4	11.5	34.6	17.3
Researchers						
SCIRE has improved my knowledge of SCI evidence	31	40.0	40.0	16.6	0	3.3
SCIRE supports my current research activities	30	26.7	56.7	10.0	3.3	3.3
SCIRE has increased my access to SCI evidence	31	43.3	33.4	13.3	6.7	3.3
SCIRE has not impacted my research activities	29	25.0	14.3	25.0	14.3	21.5

Part 4: Focus Groups

A total of 15 people over two online focus groups participated. The majority of participants were from Canada (n=10), while the remaining were from outside Canada (United States: n=3; Brazil: n=1; New Zealand: n=1). All were primarily health care providers in the field of SCI rehabilitation; however, several had some responsibilities in research (n=4) and management (n=3). The health professions represented were physical therapy (n=9), medicine (n=3), psychology (n=1), occupational therapy (n=1), and clinical educator (n=1).

From the focus groups, five major themes emerged: evidence in one place, online access, training tool, decision-making tool, and confirmatory evidence.

Evidence in One Place

A central theme emerged from a number of comments on how the SCIRE Project provided a comprehensive set of topics relevant to SCI in one place. This “one-stop shopping” was particularly advantageous when initiating a search to answer a new question: “It’s a great resource for me when I’m in a situation where I have questions about anything related to SCI, that’s probably the first place I’d start” [P4, group 2] and “If we’re doing anything new or different from the norm, they like to have a good justification. So we would often give the evidence that the SCIRE website provides” [P6, group 2].

Participants acknowledged that having the synthesized evidence could assist with the huge volume of literature: “Having a resource, you know, having this huge explosion of literature...it’s next to impossible...you’d like to keep up with everything, but it’s next to impossible to keep up with it. So, to have a resource that, like I said earlier...this is my first entry into the literature” [P4, group 2].

In addition to the evidence being in one place, the timeliness of this centralized resource was considered important: “I think it’s a great resource…and just hope that they continue to add new evidence in a timely manner” [P5, group 2] and “It’s nice to have a place that you can go that is comprehensive about current information” [P2, group 1].

Online Access

The online access was viewed as a major advantage of this knowledge tool, especially as it filled a unique niche for a health condition that had a relatively small population. The online SCIRE Project website had the advantage of being quick to access, compared to traditional literature searches. However, while there appeared to be good access to the information within specialized SCI units, some thought there was less awareness of this resource in centers that do not regularly treat SCI but could still benefit from this information:

> *First of all, the fact that it’s an online tool is excellent.* [P5, group 1]
>
> *Honestly, to my knowledge, it’s the only website like that... that exists, and being in the field of neuro and being in the field of rehab, it seems like there are a lot of resources out there for persons working with*

XSL•FO
RenderX

primarily the stroke population, but not a whole lot out there for spinal cord injuries. [P1, group 2]

The reason I don't use the publications and journal articles is sort of what the two previous people said, that when I'm on that website that's usually what I'm there for, so I'm using it fairly quickly. [P1, group 1]

I know at a rehab hospital here, we obviously know about it. But I know within the province, people in other hospitals that would treat people with SCI would probably have no idea about SCIRE. [P3, group 1]

Training Tool

Health care providers used this knowledge tool to teach others about SCI care, including staff and students. In some cases, the SCIRE Project provided the evidence to justify their current practice:

For educating new staff…it's a resource. [P2, group 1]

When I have a student and I have a SCI patient, it's a great resource to direct them to. [P5, group 2]

When I have students or I'm doing teaching, I'm on it more frequently...But it's [using SCIRE] usually when…I'm wanting to justify either my clinical practice or in my instruction to students...you want to have the evidence to back up what you're doing. [P2, group 2]

Decision-Making Tool

The SCIRE Project directly affected the clinical decision making of the health care providers, such as influencing their choice of intervention, patient equipment needs, or type of assessment tool:

A body of knowledge that a clinician can kind of depend on to help with their practice so it's definitely impacted on the way we justify things. An example, we were reassessing a C4 patient who we're looking at exercise options. They're [the physical therapy team] putting together a proposal for an FES [functional electrical stimulation] bike and we'll be using SCIRE website to help our case to provide evidence towards that. [P6, group 2]

It [SCIRE] has influenced our choice of the tool that we decided to use here for pressure ulcer prevention, so it did have a great influence on that. [P7, group 1]

I've used it clinically in terms of, for pressure ulcers…find out more about whether electrical stimulation should be used to heal pressure ulcers. So I've used that and spread it [the intervention] around to my colleagues. [P6, group 1]

I went to the website to look at the literature related to bone density and fracture risk…for clinical purposes before getting people up and trying to get them to walk. [P4, group 2]

Confirmatory Evidence

Several participants suggested that the SCIRE Project content served to confirm what they were already doing in practice and reinforced their understanding of the SCI evidence, rather than impacting their initial decision-making process:

It hasn't had a huge impact on my practice...I go there less than monthly...it's actually confirmed what I had read through other journal articles and what colleagues were saying. [P3, group 2]

It has been reinforcing in terms of if we're uncertain, in ensuring that we're doing what we should be doing. [P6, group 1]

So it [SCIRE] didn't bring any sort of new revolution with relations to practice. [P2, group 1]

Discussion

Principal Results

Taken together, the quantitative and qualitative findings of this study provide evidence that a Web-based knowledge tool may be effective in disseminating evidence-based practice, informing changes to practice, and bridging the gap between research and practice in SCI rehabilitation. The users were primarily physical therapists, physicians, occupational therapists, and nurses, which suggests that the website attracted health care providers of an interdisciplinary nature. Physical therapists were the largest group who accessed the website possibly due to the large number of interventions and outcome measures included within the SCIRE Project, which related to motor functions. Alternatively, perhaps insufficient promotion has been undertaken with other health provider groups. For example, the SCIRE Project has many interventions and outcome measures dedicated to psychosocial and community aspects of care, which would have major implications for psychologists and social workers, but these providers were underrepresented in the samples.

While the targeted end-user survey indicated that a substantial proportion (70%) were aware of the SCIRE Project resource, some indicated in the focus groups that the SCIRE Project was less well known outside of major rehabilitation or SCI centers. It is a challenge to promote the website to health care providers for whom SCI rehabilitation is a less regular part of their responsibilities (eg, family physician, community social worker). However, these providers may be most likely to benefit from the content for the small number of SCI patients they treat because they do not have the in-house support from other clinicians with SCI experience.

The focus group discussions revealed that the online aspect was an important feature that permitted fast access to the information. The online format allowed the content for a relatively specialized field to have far reach (eg, 26 countries and over 6500 users per month). The website survey and targeted end-user survey confirmed that health care providers, as well as researchers perceived that the website increased their access to SCI evidence.

The focus groups indicated that a major advantage of the SCIRE Project website was that the content was comprehensive and timely and provided one-stop shopping for the best evidence. Users could access the information when it was convenient for them. While participants accessed both the intervention and

XSL•FO
RenderX

outcome measure information, the website statistics showed a heavier access to the outcome measures. A number of studies have noted the lack of standardized measures utilized within rehabilitation [28,29]; the website provides one tool to help clinicians select appropriate outcome measures for their practice.

Impact on Practice

As intended, the SCIRE Project activities are focused on the knowledge creation domain and have a less direct role in the implementation activities of the knowledge-to-action process. However, users across all parts of the evaluation reported not only accessing the evidence but also applying the SCIRE Project evidence to their own practice contexts. Thus, access to the SCIRE Project not only improved knowledge of SCI evidence but helped inform changes to their clinical practice and improved their confidence in treating SCI clients. While all parts of the knowledge-to-action process are important, a recent systematic review showed that in the field of rehabilitation, the most common barrier to the implementation and use of evidence-based medicine was related to research utilization (eg, conflicting results, understanding statistics, literature not compiled in one place, implications for practice not being made clear) [30]. The SCIRE Project provides one important facilitator to enhancing research utilization. The SCIRE Project information directly influenced the health care provider's clinical decision making, in terms of choice of intervention, equipment needs, or assessment tool.

Some participants felt that the SCIRE Project information served more to confirm or reinforce what they were already doing in practice, rather than impacting their initial decision-making process. Researchers in other areas have noted that guidelines provide accountability in the eyes of other stakeholders [31] and may improve the clinician's confidence as it legitimizes their approach [32]. In some instances, it appeared that some health care providers accessed the SCIRE Project to justify their practice to others, especially students. The finding that users were accessing the SCIRE Project to teach others suggests a snowball effect in terms of its dissemination and impact. It is likely that to change practice, the optimal use of the SCIRE Project would be to use it in conjunction with active (eg, audit and feedback) and multicomponent implementation strategies [12,33] to facilitate uptake and change in practice. Moreover, it may be that resources like the SCIRE Project represent a platform on which to develop even more focused tools, such as clinical decision support systems that reflect customized and contextualized approaches to facilitate clinical decision making.

Limitations

As we examined the natural uptake of this website information, no control group was used to control for biases such as social desirability within the self-report responses. Furthermore, due to the fact that the samples were of convenience, they may not have been representative samples, and it is possible that those who perceived to have gained benefits from the SCIRE Project would be more likely to participate in the surveys and focus groups. The website was relatively inexpensive to develop and maintain and did not include more interactive attributes such as Web 2.0 blogs or wikis, which require more monitoring but could potentially develop a strong collaborative community of users.

Conclusions

Most notably, these findings demonstrate that the SCIRE Project was effective in disseminating information about evidence-based practice to an interdisciplinary audience and that there was some indication that it facilitated practice changes in the primary target audience of frontline clinicians. Information relating to outcome assessment represented the most frequently sought materials. It was noted that the online nature of the SCIRE Project, coupled with its comprehensiveness and timeliness, facilitated utilization as a one-stop resource for informing practice. However, there were suggestions that this resource also served to confirm or provide a rationale for existing practice, which reflects the nature of the SCIRE Project as a compilation of evidence. Additionally, there was some evidence that the SCIRE Project may have had less impact outside of major rehabilitation and SCI centers.

Taken together, these findings suggest that it might be beneficial to more explicitly describe and demonstrate to users how they might take full advantage of the SCIRE Project—not only as a knowledge creation resource (ie, evidence synthesis) but also by providing examples of how users might employ this or similar resources in their knowledge-to-action (ie, implementation) activities.

Acknowledgments

We are grateful for grant support from the Ontario Neurotrauma Foundation and the Rick Hansen Institute, in addition to salary support (to JJE) from the Canadian Institutes of Health Research (CIHR MSH-63617) and Michael Smith Foundation of Health Research.

Conflicts of Interest

None declared.

References

1. Eysenbach G. What is e-health? J Med Internet Res 2001;3(2):E20 [FREE Full text] [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
2. Cline RJ, Haynes KM. Consumer health information seeking on the Internet: the state of the art. Health Educ Res 2001 Dec;16(6):671-692 [FREE Full text] [Medline: 11780707]

XSL•FO
RenderX

3. Higgins O, Sixsmith J, Barry MM, Domegan C. A literature review on health information seeking behaviour on the web: a health consumer and health professional perspective. Stockholm: European Centre for Disease Prevention and Control; 2011.
4. Kummervold PE, Chronaki CE, Lausen B, Prokosch HU, Rasmussen J, Santana S, et al. eHealth trends in Europe 2005-2007: a population-based survey. J Med Internet Res 2008;10(4):e42 [FREE Full text] [doi: 10.2196/jmir.1023] [Medline: 19017584]
5. Maher CG, Sherrington C, Elkins M, Herbert RD, Moseley AM. Challenges for evidence-based physical therapy: accessing and interpreting high-quality evidence on therapy. Phys Ther 2004 Jul;84(7):644-654 [FREE Full text] [Medline: 15225083]
6. Younger P. Internet-based information-seeking behaviour amongst doctors and nurses: a short review of the literature. Health Info Libr J 2010 Mar;27(1):2-10. [doi: 10.1111/j.1471-1842.2010.00883.x] [Medline: 20402799]
7. Swennen MH, van der Heijden GJ, Boeije HR, van Rheenen N, Verheul FJ, van der Graaf Y, et al. Doctors' perceptions and use of evidence-based medicine: a systematic review and thematic synthesis of qualitative studies. Acad Med 2013 Sep;88(9):1384-1396. [doi: 10.1097/ACM.0b013e31829ed3cc] [Medline: 23887011]
8. Guy S, Mehta S, Leff L, Teasell R, Loh E. Anticonvulsant medication use for the management of pain following spinal cord injury: systematic review and effectiveness analysis. Spinal Cord 2014 Feb;52(2):89-96. [doi: 10.1038/sc.2013.146] [Medline: 24296804]
9. Graham ID. Knowledge synthesis and the Canadian Institutes of Health Research. Syst Rev 2012;1:6 [FREE Full text] [doi: 10.1186/2046-4053-1-6] [Medline: 22587985]
10. Straus SE, Tetroe JM, Graham ID. Knowledge translation is the use of knowledge in health care decision making. J Clin Epidemiol 2011 Jan;64(1):6-10. [doi: 10.1016/j.jclinepi.2009.08.016] [Medline: 19926445]
11. Graham ID, Logan J, Harrison MB, Straus SE, Tetroe J, Caswell W, et al. Lost in knowledge translation: time for a map? J Contin Educ Health Prof 2006;26(1):13-24. [doi: 10.1002/chp.47] [Medline: 16557505]
12. Grimshaw JM, Eccles MP, Lavis JN, Hill SJ, Squires JE. Knowledge translation of research findings. Implement Sci 2012;7:50 [FREE Full text] [doi: 10.1186/1748-5908-7-50] [Medline: 22651257]
13. Brouwers M, Stacey D, O'Connor A. Knowledge creation: synthesis, tools and products. CMAJ 2010 Feb 9;182(2):E68-E72 [FREE Full text] [doi: 10.1503/cmaj.081230] [Medline: 19884300]
14. Closs S, Lewin B. Perceptions of the barriers to research utilization: a survey of four therapies. British Journal of Therapy & Rehabilitation 1998;5(3):151-155.
15. Moore JL, Raad J, Ehrlich-Jones L, Heinemann AW. Development and use of a knowledge translation tool: the rehabilitation measures database. Arch Phys Med Rehabil 2014 Jan;95(1):197-202. [doi: 10.1016/j.apmr.2013.09.010] [Medline: 24076083]
16. Korner-Bitensky N, Roy MA, Teasell R, Kloda L, Storr C, Asseraf-Pasin L, et al. Creation and pilot testing of StrokEngine: a stroke rehabilitation intervention website for clinicians and families. J Rehabil Med 2008 May;40(5):329-333 [FREE Full text] [doi: 10.2340/16501977-0177] [Medline: 18461256]
17. Browman GP. Scientific evidence and expert opinion in ACP Journal Club: commentary on commentaries. ACP J Club 1994;120 Suppl 1:A11. [Medline: 8137125]
18. Grol R. Successes and failures in the implementation of evidence-based guidelines for clinical practice. Med Care 2001 Aug;39(8 Suppl 2):II46-II54. [Medline: 11583121]
19. Morris ZS, Wooding S, Grant J. The answer is 17 years, what is the question: understanding time lags in translational research. J R Soc Med 2011 Dec;104(12):510-520 [FREE Full text] [doi: 10.1258/jrsm.2011.110180] [Medline: 22179294]
20. Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ, Montori VM. Internet-based learning in the health professions: a meta-analysis. JAMA 2008 Sep 10;300(10):1181-1196. [doi: 10.1001/jama.300.10.1181] [Medline: 18780847]
21. Eng JJ, Teasell R, Miller WC, Wolfe DL, Townson AF, Aubut JA, the SCIRE Research Team. Spinal Cord Injury Rehabilitation Evidence: Methods of the SCIRE Systematic Review. Top Spinal Cord Inj Rehabil 2007;13(1):1-10 [FREE Full text] [doi: 10.1310/sci1301-1] [Medline: 22767989]
22. Craven C, Balioussis C, Verrier MC, Hsieh JT, Cherban E, Rasheed A, et al. Using scoping review methods to describe current capacity and prescribe change in Canadian SCI rehabilitation service delivery. J Spinal Cord Med 2012 Sep;35(5):392-399 [FREE Full text] [doi: 10.1179/2045772312Y.0000000045] [Medline: 23031176]
23. Pannek J. Treatment of urinary tract infection in persons with spinal cord injury: guidelines, evidence, and clinical practice. A questionnaire-based survey and review of the literature. J Spinal Cord Med 2011;34(1):11-15 [FREE Full text] [doi: 10.1179/107902610X12886261091839] [Medline: 21528621]
24. Thomason SS, Evitt CP, Harrow JJ, Love L, Moore DH, Mullins MA, et al. Providers' perceptions of spinal cord injury pressure ulcer guidelines. J Spinal Cord Med 2007;30(2):117-126 [FREE Full text] [Medline: 17591223]
25. Pathman DE, Konrad TR, Freed GL, Freeman VA, Koch GG. The awareness-to-adherence model of the steps to clinical guideline compliance. The case of pediatric vaccine recommendations. Med Care 1996 Sep;34(9):873-889. [Medline: 8792778]
26. Dillman DA, Smyth JD, Christian LM. Internet, mail, and mixed-mode surveys: the tailored design method. Hoboken, NJ: Wiley & Sons; 2009.
27. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today 2004 Feb;24(2):105-112. [doi: 10.1016/j.nedt.2003.10.001] [Medline: 14769454]

XSL•FO
RenderX

28. Swinkels RA, van Peppen RP, Wittink H, Custers JW, Beurskens AJ. Current use and barriers and facilitators for implementation of standardised measures in physical therapy in the Netherlands. BMC Musculoskelet Disord 2011;12:106 [FREE Full text] [doi: 10.1186/1471-2474-12-106] [Medline: 21600045]
29. Ali M, English C, Bernhardt J, Sunnerhagen KS, Brady M, VISTA-Rehab Collaboration. More outcomes than trials: a call for consistent data collection across stroke rehabilitation trials. Int J Stroke 2013 Jan;8(1):18-24. [doi: 10.1111/j.1747-4949.2012.00973.x] [Medline: 23280265]
30. Sadeghi-Bazargani H, Tabrizi JS, Azami-Aghdash S. Barriers to evidence-based medicine: a systematic review. J Eval Clin Pract 2014 Aug 18:-. [doi: 10.1111/jep.12222] [Medline: 25130323]
31. Van der Wees PJ, Moore AP, Powers CM, Stewart A, Nijhuis-van der Sanden MW, de Bie RA. Development of clinical guidelines in physical therapy: perspective for international collaboration. Phys Ther 2011 Oct;91(10):1551-1563 [FREE Full text] [doi: 10.2522/ptj.20100305] [Medline: 21799137]
32. Poitras S, Durand MJ, Côté A, Tousignant M. Use of low-back pain guidelines by occupational therapists: a qualitative study of barriers and facilitators. Work 2011;39(4):465-475. [doi: 10.3233/WOR-2011-1196] [Medline: 21811035]
33. van der Wees PJ, Jamtvedt G, Rebbeck T, de Bie RA, Dekker J, Hendriks EJ. Multifaceted strategies may increase implementation of physiotherapy clinical guidelines: a systematic review. Aust J Physiother 2008;54(4):233-241. [Medline: 19025503]

Abbreviations

SCIRE: Spinal Cord Injury Rehabilitation Evidence
SCI: Spinal cord injury

Edited by G Eysenbach; submitted 04.04.14; peer-reviewed by A Thomas, I David; comments to author 16.07.14; accepted 23.11.14; published 23.12.14.

Please cite as:
Eng JJ, Noonan VK, Townson AF, Higgins CE, Rogers J, Wolfe DL
Impact of an Online Medical Internet Site on Knowledge and Practice of Health Care Providers: A Mixed Methods Study of the Spinal Cord Injury Rehabilitation Evidence Project
J Med Internet Res 2014;16(12):e296
URL: http://www.jmir.org/2014/12/e296/
doi:10.2196/jmir.3453
PMID:25537167

©Janice J Eng, Vanessa K Noonan, Andrea F Townson, Caroline E Higgins, Jess Rogers, Dalton L Wolfe. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 23.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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

Occupational Therapists' Views on Using a Virtual Reality Interior Design Application Within the Pre-Discharge Home Visit Process

Anita Atwal[1], OT, DipCOT, MSc, PhD; Arthur Money[2], BSc (Hons), MSc, PhD Comp Sc; Michele Harvey[1], BSc (Hons), MSc

[1]Brunel University, School of Health Sciences and Social Care, Brunel University, London, United Kingdom

[2]Brunel University, Department of Computer Science, Brunel University, London, United Kingdom

Corresponding Author:
Arthur Money, BSc (Hons), MSc, PhD Comp Sc
Brunel University
Department of Computer Science
Brunel University
St Johns
London, UB8 3PH
United Kingdom
Phone: 44 1895266758
Fax: 44 1895266758
Email: arthur.money@brunel.ac.uk

Abstract

Background: A key role of Occupational Therapists (OTs) is to carry out pre-discharge home visits (PHV) and propose appropriate adaptations to the home environment in order to enable patients to function independently after hospital discharge. However, research shows that more than 50% of specialist equipment installed as part of home adaptations is not used by patients. A key reason for this is that decisions about home adaptations are often made without adequate collaboration and consultation with the patient. Consequently, there is an urgent need to seek out new and innovative uses of technology to facilitate patient/practitioner collaboration, engagement, and shared decision making in the PHV process. Virtual reality interior design applications (VRIDAs) primarily allow users to simulate the home environment and visualize changes prior to implementing them. Customized VRIDAs, which also model specialist occupational therapy equipment, could become a valuable tool to facilitate improved patient/practitioner collaboration, if developed effectively and integrated into the PHV process.

Objective: The intent of the study was to explore the perceptions of OTs with regard to using VRIDAs as an assistive tool within the PHV process.

Methods: Task-oriented interactive usability sessions, utilizing the think-aloud protocol and subsequent semi-structured interviews were carried out with seven OTs who possessed significant experience across a range of clinical settings. Template analysis was carried out on the think-aloud and interview data. Analysis was both inductive and driven by theory, centering around the parameters that impact upon the acceptance, adoption, and use of this technology in practice as indicated by the Technology Acceptance Model (TAM).

Results: OTs' perceptions were identified relating to three core themes: (1) perceived usefulness (PU), (2) perceived ease of use (PEoU), and (3) actual use (AU). Regarding PU, OTs believed VRIDAs had promising potential to increase understanding, enrich communication and patient involvement, and improve patient/practitioner shared understanding. However, it was unlikely that VRIDAs would be suitable for use with cognitively impaired patients. For PEoU, all OTs were able to use the software and complete the tasks successfully; however, participants noted numerous specialist equipment items that could be added to the furniture library. AU perceptions were positive regarding use of the application across a range of clinical settings including children/young adults, long-term conditions, neurology, older adults, and social services. However, some "fine tuning" may be necessary if the application is to be optimally used in practice.

Conclusions: Participants perceived the use of VRIDAs in practice would enhance levels of patient/practitioner collaboration and provide a much needed mechanism via which patients are empowered to become more equal partners in decisions made about their care. Further research is needed to explore patient perceptions of VRIDAs, to make necessary customizations accordingly, and to explore deployment of the application in a collaborative patient/practitioner-based context.

XSL•FO
RenderX

(J Med Internet Res 2014;16(12):e283) doi:10.2196/jmir.3723

KEYWORDS

occupational therapy; pre-discharge home visits; virtual reality; 3D; patient collaboration; patient engagement; empowerment; technology assisted care; patient perceptions

Introduction

Background

With an anticipated rise in the demand for health care resources as a result of an ageing population [1], government initiatives see innovations in technology for health care as one of the few areas in which there still remains capacity for reducing costs and improving quality of service for patients [2]. In particular, the recently proposed long term version for the National Health Service (NHS) under the banner of "Equity and Excellence: Liberating the NHS" [3] demonstrates the UK government's commitment to innovation via the use of Information and Communication Technology (ICT). ICT is seen as a key lever in delivering person-centered, preventative, re-abling, and personalized care. If this vision is to be realized, it is crucial to "empower and liberate clinicians to innovate"' [3], hence enabling practitioners to adopt and integrate new technologies and practices with a view to improving patient health outcomes. A central role of incorporating ICT into heath care delivery is to provide more effective patient–centred health care that creates opportunities for both practioners and patients to work collaboratively in the consultation and decision making processes. [4,5]. Enabling patient/practitioner collaboration will improve the extent to which the patient is aware of their health issues, consequently improving levels of patient engagement, adherence, and satisfaction [6]. Promoting innovative applications of technology for health care is seen as playing a central role in enabling patients to take responsibility for their own care, and improve and sustain quality of life by making it possible to live independently within their own homes for longer [7].

A primary area of focus within the domain of Occupational Therapy is to enable patients to live independently within their homes. In order to facilitate this, a key role of an occupational therapist (OT), across Europe, Australia, and North America [8], is to carry out a pre-discharge home visit (PHV) with the patient to facilitate appropriate, safe, and successful discharge from hospital to home [9]. The aim of the PHV is for the clinician to visit the home with the patient to provide additional information about how the patient will cope within their home environment after discharge. During the visit the OT may propose modifications to the home environment, where appropriate, to enable the patient to function at a satisfactory level of independence after discharge [10]. Modifications to the home environment may include installation of specialist assistive equipment such as bed hoists, support rails, shower seats, grab rails, and so forth. Furthermore, often patients may be faced with the prospect of using a wheelchair or walking frame to aid mobility, which may also dictate that alterations must be made to the layout of the home to accommodate access. Recent research in the field of occupational therapy has revealed that PHVs can be cumbersome, highly resource intensive, and sub-optimal for patients [8,11]. One significant issue is that decisions made regarding home adaptations are often made without adequate collaboration and consultation with the patient [10,12]. Currently, the only real opportunity to consult and collaborate with the patient is while the clinician and patient are together in the home as part of the PHV process. However, patients have reported that they find this anxiety-provoking and feel as if they are being tested/assessed in terms of their mobility around the home; hence, they do not feel able to collaborate as an equal partner for fear of not being discharged home after the visit [10]. This could influence use of assistive devises since more that 50% of home adaptations are not used after discharge home, resulting in sub-optimal health outcomes and significant wastage of resources [12]. This is perhaps no surprise when considering how personalized and sensitive the home setting is [13], Coupled with the fact that there is no readily available tool or technique that assists patients and practioners to jointly visualize and explore the home environment according to the patients' personal needs [14].

There is a need to seek out and develop new and innovative uses of technology that enable patients and practitioners to jointly understand and visualize the complexities and meanings associated with the home environment, to envisage the challenges that are likely to be encountered within the home, and to collaborate and contribute equally to developing solutions to these challenges [15]. This is likely to lead to many positive health outcomes such as improved adherence, engagement, and patient satisfaction [6].

Virtual Reality Interior Design Applications for Occupational Therapy

Over the past decade, Virtual Reality (VR) has become a valuable tool that has been applied to a range of health care scenarios [16]. The term VR typically relates to interactive three-dimensional (3D) computer-generated environments that simulate being present within the real world equivalent of that environment [17]. Application areas of VR span across a range of domains including interior design, health care, military and defense, education, and entertainment and gaming [18]. Specifically relating to health care, perhaps the most well-noted application of VR has been for the treatment of phobia such as public speaking anxiety [19] and the treatment of claustrophobia [20]. More specifically within the domain of home interior design, virtual reality interior design applications (VRIDA) serve as a valuable assistive tool for negotiating adaptations between designers and home owners [21]. VRIDA allows individuals to design or redesign their homes virtually, prior to making these changes a reality. The advantages of using VRIDA in interior design include improved collaboration between the home owner and the designer, and enhanced understanding and communication of design options. It also brings design misconceptions to the forefront of discussion, facilitating active participation by all parties involved, and aiding the process of

XSL•FO
RenderX

achieving consensus between all parties [22,23]. Figure 1 provides some examples of 3D home environments produced using VRIDA. The specific VRIDA used to produce the example environments was "SweetHome 3D", a freely available open-source 3D interior design software application [24].

In light of the need for improved collaboration between OTs and patients, this research proposes to explore the use of VRIDA to aid the PHV process and gain insights into patient and practitioner experiences of its application in practice. The prospect of using VRIDA has potential to respond to a number of the issues that currently limit the effectiveness of PHVs. VRIDA would serve as a tool that enables occupational therapists to rapidly create the 3D representation of the patient's home, allowing the patient and practitioner to jointly visualize the interior of the home and trial a range of adaptations and specialist equipment within it. This would enhance collaboration between clinician and patient and assist them in making shared decisions about how this sensitive and personalized space may be best adapted specifically to the patient's individual needs. It would also provide an interactive simulation of the home, enabling the patient to "walk" through the home, via a personal computer or laptop, which could help therapists to better consider barriers to everyday performance and enhance the patient's insight and motivation to participate in tailored interventions. VRIDA would provide the patient with the valuable opportunity to consult as an expert on their own needs, and participate as an equal partner in decision making, without feeling as if their mobility is being assessed, as is often the case when visiting the home in person with the practitioner. To date, however, the gains that VRIDA could bring to occupational therapy practice are yet to be capitalized upon, as little research has been carried out within this particular health care context.

Figure 1. Examples of virtual home environments, lounge (left), kitchen (right), produced using Virtual Reality Interior Design Applications (VRIDA).



Practitioner Perceptions and Technology Acceptance

The insight of practitioners is extremely valuable and should be employed at all stages of technology development and deployment. Research with health care practitioners has shown that they are more likely to adopt technologies if these are viewed as compatible with current practice [25]. Compatibility issues could relate to time, existence of evidence on positive outcomes in practice [26], organizational issues [27], and attitudes of professionals towards technology [28]. Therefore, the application of VRIDA must be accepted by therapists if it is to serve as a feasible tool that may be used within occupational therapy in practice [28]. If VR therapies are not perceived as usable or likable in actual clinical settings, it is unlikely that the technology will remain in use long enough for an evidence-base to be explored and established.

Over the past two decades, much research effort has been invested into understanding end users' reactions and motivations to technology acceptance, adoption, and use [29]. Although a large portion of this research effort has been focused within the Information Systems research domain, more recently, there is increasing interest in gaining a better understanding of the factors that influence user acceptance, adoption, and use of technology within the health care domain [30]. To the best of our knowledge, there is no research yet that explores barriers to acceptance, adoption, and use of VRIDA technology for use within the occupational therapy context or its application to the PHV process. The Technology Acceptance Model (TAM) is perhaps the most notable theory applied in the explanation of user motivations, attitudes, and responses to acceptance and use of technology [31]. Despite its relative simplicity, even the most basic form of TAM is typically seen to provide an explanation of approximately 40% of issues related to technology acceptance [29].

TAM proposes that when presented with a new technology, users' behavioral intention to use and their Actual Use (AU) of technology are typically mediated by two key factors: Perceived Usefulness (PU), which is the extent to which the user perceives that the new technology will aid them in performing the task at hand, and Perceived Ease of Use (PEOU), which is the extent to which the individual believes using the technology would be free of effort [32]. TAM is now increasingly being applied within the health care research domain [33]. Examples include exploring the acceptance of: telemedicine technology by nurses [34], Personal Digital Assistants (PDAs) by physicians [35], portable postural assessment technologies for use by physiotherapists [36], mobile picture archiving technologies for dental care [37], and a range of customizable and wearable health care devices for patients and practitioners [38]. Although the vast majority of TAM research to date has been quantitative, there is increasing recognition that qualitative enquiry, particularly in the early explorative stages, is well suited to

XSL•FO
RenderX

scoping the design and development of new innovations and identifying the range of factors that may affect the acceptance, adoption, and use of a specific technology [39].

The aim of this study is to explore occupational therapists' perceptions of VRIDA and to gain insights into the feasibility of using VRIDA as a tool to aid the PHV process in relation to the key factors outlined in the technology acceptance model. The next section provides details of the study carried out to achieve this aim. The results of this study are then presented, followed by a discussion of the implications of the findings in the context of existing research literature and outlining the study limitations. Finally, the study is concluded and future research directions are considered in light of the findings.

Methods

Overview

The aim of this study was to explore the perceptions of OTs relating to the three TAM factors (PU, PEOU, and AU) and the potential feasibility of using VRIDA applications as an assistive tool that may be used within the PHV process.

Participants

A convenience sampling strategy was used for recruitment of participants for this study. The inclusion criteria were that participants were practicing OTs within the UK health sector and that they were familiar with using desktop computers and typical applications such as Microsoft Word and accessing email. Potential participants were primarily identified from the researchers' existing social network contacts list (ie, LinkedIn contacts) and subsequently contacted by email in the first instance and invited to take part in this study. No financial incentives were offered to take part in the study, hence participation took place purely on a voluntary basis. A total of seven OTs were recruited and took part in the study. This number of participants is in excess of the recommended threshold of five participants typically required to carry out effective think-aloud interaction and usability testing [40,41]. Five of the participants were female and two were male. The amount of clinical experience that participants had ranged from 5-20 years and areas of specialty included community-based social services, older persons, mental health, acute care, and pediatrics. Table 1 provides a summary of the participant profiles.

Table 1. Summary of participant profiles.

Participant	Gender	Years practicing	Area of specialty
A	Male	More than 5 years	Social services (community)
B	Female	More than 10 years	Senior Therapist Older Person
C	Female	More than 5 years	Senior Therapist Older Person
D	Female	Less than 5 years	Social services (community)
E	Male	More than 20 years	Mental Health Team Leader
F	Female	More than 10 years	Senior Therapist Acute Care
G	Male	More than 5 years	Pediatrics

Orientation Task Using a VRIDA

On arrival, information sheets were distributed to users prior to participation in the session, the content of which was worked through with each participant individually. The information sheet provided a brief background and context and purpose to the study, and summarized the main activities that would take place during the course of the session. Participants were encouraged to ask questions throughout the process, and any questions were answered as they arose. Participants were then asked to complete a consent form in which their ethical rights were explained in terms of informed consent, withdrawal, and anonymity.

Participants were given the task of using a VRIDA to design the interior of a room that would typically represent a patient's home environment. The VRIDA software application used for the purposes of this task was a customized version of SweetHome 3D [24]. Figure 2 illustrates the SweetHome 3D application interface used by participants to design and develop home interiors.

The SweetHome 3D application interface is made up of four main functional quadrants: (1) furniture catalogue, (2) home plan, (3) home furniture list, and (4) 3D view. For the purposes of this study, the application has been customized to include a library of specialist OT assistive equipment necessary for OTs to make typical home adaptation recommendations as part of the PHV process. These artefacts were presented within the furniture catalogue quadrant of the application in a folder entitled "OT Objects". Occupational therapy assistive devices featured in the library included ramps, a range of grab rails, a bath hoist, a wheelchair, toilet frame, and seat. The custom OT objects library folder, how this was integrated into the furniture catalogue navigation pane, and examples of some of these OT objects (wheelchair and toilet frame) are presented in Figure 2, as well as how these OT objects may be modelled within an example 3D view of a bathroom environment.

Prior to the main task of designing a typical patient room of their choice, participants were provided with basic written instructions, presented in Table 2, outlining the key steps necessary to create a room using the application.

Printed screenshots of the SweetHome 3D interface and the 4-quadrant map of the software (similar to those presented in Figure 2) were also provided alongside the written instructions. Using the resources provided, participants were asked to design

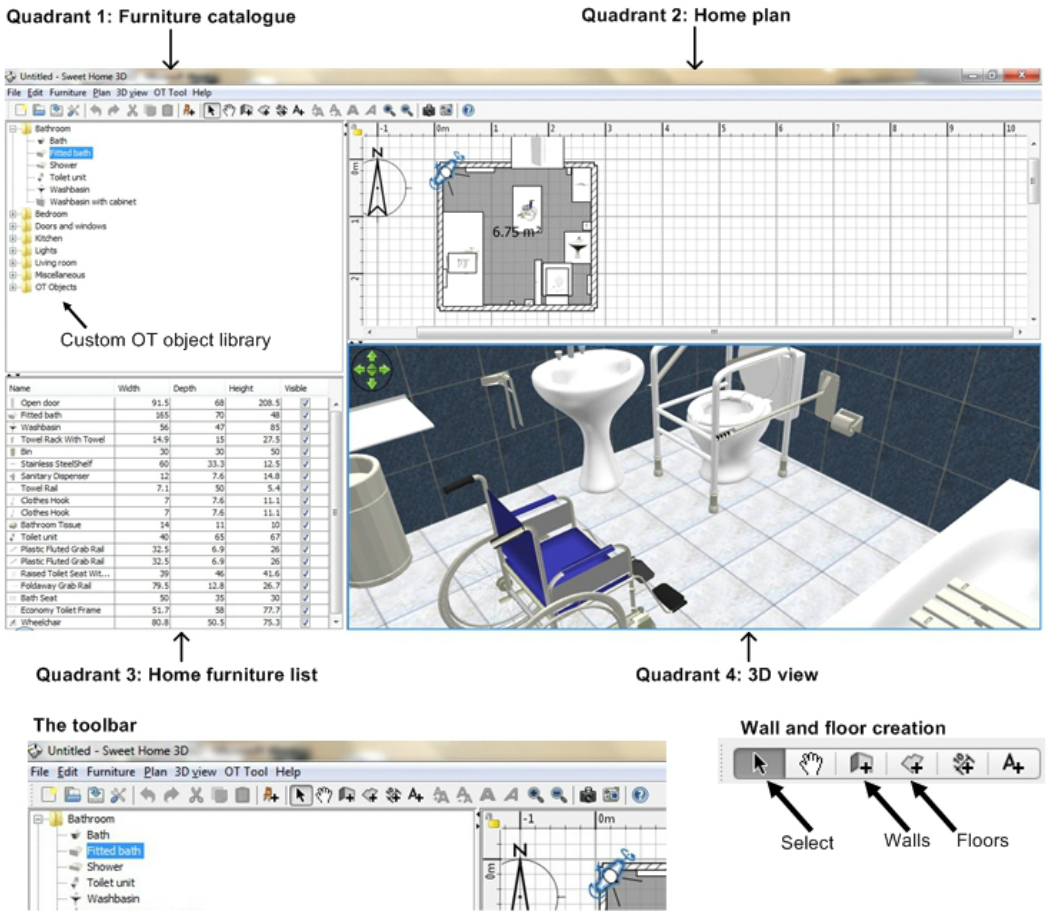
XSL•FO
RenderX

a bathroom in order to practice developing an environment prior to moving on to the main task. After the participants perceived they were confident in utilizing the software, they were asked to proceed with the main task.

Table 2. Written instructions for initial familiarization and orientation with SweetHome 3D.

Instructions	
Create your room	
	Draw floor in Quadrant 3 using the Floor button (follow instructions in pop-up box)
	Draw walls in Quadrant 3 using the Walls button (follow instructions in pop-up box)
Furnish your room	
	Choose objects from Quadrant 1 using the Select tool
	Drag and drop into Quadrant 3, arrange using the Select tool
	An inventory of these objects will appear in Quadrant 2
Decorate your room	
	Select a wall or floor in Quadrant 3, it will highlight in blue once it is selected
	Right-click on highlighted wall or floor, select “Modify Walls” or “Modify Floor”, choose colors/textures
Visit your room	
	Go to “3D View” menu at the top menu, choose “Virtual Visit”
	A figure will appear in Quadrant 3 and the view in Quadrant 4 will change
	Move and click in Quadrant 4 to look around the room
Save your room	
	Go to “File” at the top menu
	Name your file and save to the desktop

Figure 2. The customized SweetHome 3D application interface.



XSL•FO
RenderX

Think-Aloud Interior Home Design Task

For the main task, participants were asked to design a room of their choice, which they believed represented a typical room within a patient's home. They were also welcome to insert assistive equipment where they deemed necessary. Participants were asked to design the room of choice from scratch, while adopting a 'think-aloud' approach, which enabled them to verbally share their thoughts while interacting with the application [42]. Think-aloud is a valuable research technique that is typically used for the real-time capture of, not only client preferences and thoughts while interacting with software applications, but also the reasoning behind these preferences and thoughts. It is frequently used within usability testing and has been used within occupational therapy to explore clinical reasoning [43]. It was important that users did not feel pressured, as it was anticipated that this may impact on the level of think-aloud thoughts they felt comfortable sharing during the task. Therefore, users were reminded throughout the task that there was no urgency in completing the task, and encouraged to take as long as they required to provide comments and to interact with the SweetHome 3D application. During the sessions, standard prompts to think aloud, such as "what are you thinking?" and "what are you doing now?" were used whenever the researcher felt there were extended periods of silence. The use of task-focused prompts such as these ensured that participants' attention remained on the task at hand but also that sufficient data and commentary relating to their interactions with the software application were elicited [44]. All sessions were audio-recorded and the researcher also took written notes during the session. At the end of the session, a discussion was held with the user, giving them the opportunity to elaborate on any of the points they made during the session and to reflect on their experience of using the application and perceptions relating to its ease of use, usefulness, usability, and the feasibility of it actually being used as a tool to assist in the PHV process.

Data Analysis

Template analysis was used to analyze the interview data. This is a form of thematic analysis, which involves development of a coding template that represents a summary of the themes that are seen by the researcher(s) as being of importance within the dataset [45]. The approach taken to this was both inductive, as some themes were closely linked to the data, and other themes were driven by theory and the researchers' analytical interest [46]. Analysis often begins with some *a priori* themes/codes that are of interest to the researcher. In this case, the interview data was approached with the broad aim of exploring factors that are related to TAM. Hence, analysis considered the participant perceptions of the feasibility, usability, and utility of VRIDA for PHV interventions within the context of three high-level TAM themes relating to PU, PEOU, and AU. The approach taken in the analysis of this data is in line with what Madhill et al [47] refer to as the 'contextual constructivist' position. In this case, it is accepted that there are many interpretations that may be made of a given phenomenon, which depends upon the focus of the researcher and the context in which the research is carried out. Hence, the themes and sub-themes that emerge as a result of the analysis are partly a product of these factors.

As an initial step, all interview recordings were transcribed into text format. The textual dataset in its entirety was perused to conceptualize the data and its relationship to the *a priori* themes that existed at a high level. The entire dataset was closely read and patterns in the data were noted. Sequences of data that represented "the most basic segment, or element, of the raw data or information that can be assessed in a meaningful way regarding the phenomenon" [48] were identified and assigned a code name. The dataset was then examined iteratively through several stages of splicing, linking, deleting, and reassigning themes and sub-themes. In this way, a final representation of the themes in the data was produced. The themes did not 'emerge', because they did not have a concrete existence in the data, rather they were constructed as part of the interpretative work. The analysis goes beyond the surface meaning of the data and tries to "identify the underlying ideas, assumptions, and conceptualizations—and ideologies—that are theorized as shaping or informing the semantic content of the data" [46]. The first and second authors coded the data and discussed inconsistencies where these arose until a clear consensus of the main themes was reached. The main themes are those drawn from multiple contributions and that represent issues that are clearly central to the participants themselves. Within these themes, we have identified sub-themes that depict the breadth of positions that were adopted within the main themes. For a detailed description of the thematic template analysis process, see King [45], and similarly for the thematic analysis process, see Joffe and Yardley [49] and Silverman [50].

Ethical Considerations

The study was reviewed and approved by the Brunel University Research Ethics Committee prior to any data collection. All participants taking part in the study were guaranteed confidentiality and anonymity. Signed consent forms were obtained from all participants prior to taking part in the semi-structured interviews. Participants were informed of their right to withdraw from the study at any time. This was done both in writing and verbally.

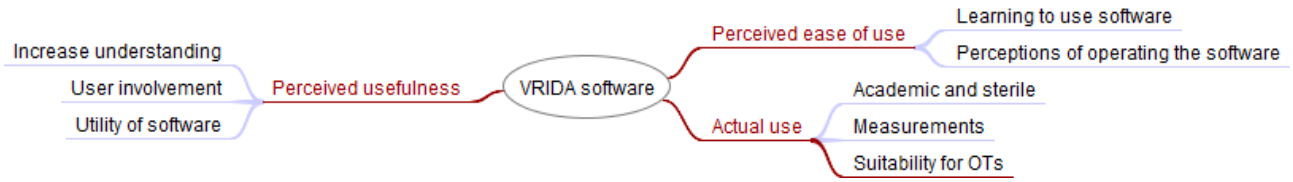
Results

Overview

The results of the analysis of think-aloud responses and the discussions held at the end of each session are presented in this section in the context of the three key TAM themes used for analysis: PU, PEOU, and AU. A number of sub-themes were identified within these key TAM themes, these are presented as a thematic mind map in Figure 3.

XSL•FO
RenderX

Figure 3. Thematic mind map of themes and sub-themes.



Perceived Usefulness (PU)

Increased Understanding

Participants B, D, E, F, and G felt that 3D images were a good visual aid that enabled patients to have a better understanding of assistive technology or adaptations to be provided. It was felt that the rich visual representations and interactive environment provided by the application is preferable to the static hand-drawn examples that are often used in practice. The 3D images were also seen to somehow convey additional information that otherwise would be difficult to verbalize in their absence.

> *I think everyone could have a look and it's much better than having drawings or trying to explain them.* [Participant D]

Other participants perceived that it can give the patient "immediate feedback" on planned changes (Participants C, G), which is likely to improve shared understanding of proposed adaptations to the home and the extent to which patients and practitioners can engage in a meaningful discussion about a particular scenario. Participant A felt that comprehension and evaluation of items that might be difficult to determine when looking at an aerial or 2D drawing would be much clearer when presented as a 3D representation. This participant also felt that it would be helpful to spot additional issues that would be lost using 2D representations, such as the height of an oven for a wheelchair user.

User Involvement

Participants A, E, F, G spoke about the positive impact of client involvement. They suggested that the use of the application in conjunction with the patient would be likely to empower patients and enable them to share the expertise and knowledge that they have about their unique circumstances, how they manage their condition, and how they engage with their home environment.

> *They are the experts in their situation and so if we can get them to join in with the design process, it makes it easier for everybody.* [Participant A]

The software was perceived as supporting shared decision making since patients would be more involved within the process (B, E) and understand the rationale behind the suggested adaptations. Hence, participants felt that the visualization afforded by the application would help to foster improved levels of engagement in the PHV process and reduce the levels of ambivalence and anxiety that sometimes surround the process of introducing assistive equipment into the home environment.

> *Most people are really a bit ambivalent about equipment. Obviously, first it's a horrible reminder of things going wrong. And usually it's because a lot of people can't visualize it.* [Participant E]

The application was also perceived as a tool that could be used for the independent assessment of technology and enable parents to design environments for their children (Participant A, C, G). In particular, the negotiation process for introducing new equipment into the home can sometimes be extremely resource intensive, requiring numerous visits to the home to explore concerns regarding space requirements and positioning of equipment. Utilization of the VRIDA application was seen to offer a solution that could potentially reduce the time and resources required to come to an agreement on home adaptations.

> *Because every parent says they don't have the space, so when you can show you have room for that, it would be so different what you can do without going so many times in his house* [Participant G]

Utility of Software

One perception (Participant D) was that the software would not be suitable for persons with cognitive impairment. It was felt that such patients may find it challenging to make the connection between the virtual representation of the home environment and the home itself. However, some participants believed that the application would be very helpful when carrying out major modifications to the home environment (Participant B, E) and was more effective than the current method of taking photographs (Participant C, D) or paper drawings (Participant A). One of the advantages of the interactive 3D representations would be that the patients could immediately see the proposed changes, without having to worry about the quality or scaling of the photographs or hand drawings that have been used to come to a decision. It was felt that 3D representations were likely to offer peace of mind and foster better quality and more timely collaboration around a more accurate representation of the individual patient's home setting.

> *If you can knock something like this up and send it to them, and even better these days when so many people are online, you can create something back at the office and send it through to them and say, right, this is what I think, what do you reckon?* [Participant A]

Perceived Ease of Use (PEOU)

Learning to Use the Software

All of the participants were able to complete the assigned tasks and vocalized the way they best learned the new tasks. Participants A, F, E, and G stated that they did not like or use the written instructions provided to assist them in learning how to use the application. All reported that they were able to make sense of the key application functions fairly intuitively without

XSL•FO
RenderX

any assistance. Both participants F and G explicitly stated that they preferred to learn by playing with the software as opposed to following written instructions. In contrast, participants B, C and D favored the use of the written instructions for guidance and liked their conciseness and the narrative, which enabled them to engage in a simple task before moving on to designing a home environment of their own conception. Only one participant (Participant C) emphasized the overhead of effort required to practice utilizing the software before any significant progress was made.

Operating the Software

A number of issues relating to its usability were identified by participants as a result of carrying out the main task. Participants A, B, and F all commented on how they had difficulty picking up or selecting an item of furniture in order to move it to a new position within the room while completing the task. Participants B, D, E, and F stated that the controls were sometimes tricky to operate. Participants B and G felt that the rotation of objects, in particular, should be done with a dedicated button that would move the object in 90 degree turns, similar to how photos are rotated in digital photo viewing applications. Participants A, B, C, D, E, and F experienced issues or confusion over how to apply a wall texture or color; specifically, the software terms of "left side" and "right side" to wall orientation were unclear. Similarly, Participants B, C, and F felt that the default white color for the floors, walls, and objects made it difficult to visually differentiate between them. Participants B, E, and F felt that the mouse controls were too sensitive.

Participants also made numerous suggestions about additional items that should be included in future versions of the furniture library and OT object catalogue. A summary of the additional items of furniture and assistive equipment suggested by participants are presented in Table 3.

Table 3. Suggested additional items for OT[a]object library.

Participant(s)	Suggested item(s)
A, E	Ceiling track hoist
A	Drain (shower room)
F	Folding door or "doors that go both ways"
G	Mirror
C	Non-slip mat
D	OT items for bedroom
C	OT items for kitchen
G	OT items for children's playroom
A, C, D, E, F	Rails in multiple lengths/rotations
A	Ramps (outdoor items)
E	Sash window
A	Wall-hung basin
A, B	Wheelchair turning radius graphic

[a]OT: occupational therapy

Actual Use (AU) of the Technology

Overview

In general, most participants were positive about the use of the software. Some of the comments were: "it's quite cool", "My kids would love it", and "it's really great". Some participants were positive about the value that this application could deliver to occupational therapy more generally and across a range of services.

> *I think it's doing a great job. I was really impressed. I think that can really help many OTs throughout the country.* [Participant D]

Two participants perceived that more work needed to be done if maximum benefits of the application were to be realized in practice. However, they noted that the majority of functionality currently offered is useful and that with only minor adaptations to the interface and functionality, the application would be beneficial to use in practice.

> *I'm sure it's a case of fine tuning rather than significant changes.* [Participant A]

Look and Feel

Participants B and G felt that the look and feel of the digital home images needed enhancement to enable a client to connect with the 3D images more effectively. They felt that the home environments presented within the application felt slightly artificial in some way and could benefit from being softened or made to look more 'lived-in'.

> *I think it's got potential, but it still feels quite academic, quite sterile.*

Some participants made suggestions relating to how the modelled environments could be made to feel more life-like and lived-in. For example, both Participants B and G suggested making simple additions, such as a rubber duck in the bathroom,

XSL•FO
RenderX

to help to add a home-like element that was otherwise felt to be missing. Participant B also suggested a towel on the towel rail, bottles of shampoo, blinds/curtains on the windows, and houseplants.

> *If you have a bath, where is the bottle of shampoo? Because that's going to make it look like it is someone's home. It is easier, it opens up the ability to engage with people who maybe need that household.* [Participant B]

Participants B and F felt that this technology, specifically the use of a computer and mouse, was outdated. They felt that perhaps delivering interventions using a VRIDA application may be better delivered on more mobile types of platform, such as a tablet computer or a laptop.

Measurements

Participants emphasized the importance of measurement and having objects in pre-set sizes (Participants A, C). They felt that it is important to ensure that assistive pieces of equipment are modelled to scale within the environment.

> *Because at the moment you could end up with a design that looks wonderful but you can't actually achieve it because you have dropped in a bath that is not actually on the market.* [Participant A]

Indeed, Participants B and D thought that Sweet Home 3D with exact measurements could be a beneficial tool when communicating with assistive equipment installation technicians. They also felt the exact measurements in Sweet Home 3D would give clients a better representation of what they would be receiving and how it would be oriented. Participant A felt that standardized objects should be included and that the ability to resize or stretch objects in Sweet Home 3D may lead to errors, therefore, suggesting that the sizes of objects within the OT object library should be fixed and protected against being resized within the application.

Suitability for OTs

There was a view that the software may be suitable for actual use by OTs working with a variety of clinical conditions. Table 4 summarizes the types of clinical conditions for which SweetHome 3D is likely to be usefully applied to.

One participant perceived that if the technology improved, they would "take it on all of my visits" (Participant G). Another participant perceived that if the technology was to be used in practice, then it must be used with a tablet computer (Participant B).

However, one participant worried about the impact on the profession and was concerned that introduction of such technology could potentially result in less OTs being employed within the profession.

> *But the one thing I would say, this isn't necessarily criticism, but it's just whether not – because it's been so easily done, whether that actually de-skills OTs and actually kind of takes their jobs away. You won't need them anymore because you've got a whiz-bang computer that can do it for you.* [Participant E]

Table 4. Suggested clinical areas.

Participant(s)	Suggested clinical usage
B, F, G	Children/Young Adults
B	Clients that are difficult to engage
B, D	Long-term Conditions
F	Neurology
F	Older Adults
D	Social Services

Discussion

Principal Findings

In this study, occupational therapists viewed the VRIDA software as being a potentially important and useful visual aid to facilitate shared understanding and shared decision making about home adaptations with patients. This is particularly valuable given that, to date, insufficient explanation and notification of home adaptations during home visits has resulted in some users feeling dissatisfied with their experience, resulting in equipment abandonment levels in excess of 50% [12,51,52]. Enabling people to stay at home and maintain independence at home can add to an increased sense of control and improved quality of life [53-55]. The OTs that took part in this study did not seem to be concerned about the usability of the software but rather were more interested in the impact of using a VRIDA as a tool to assist in occupational therapy interventions and the positive impact this would have on the patient experience. Interestingly, evidence from a study involving older adults from 11 European countries found that older adults wanted to have a trusting relationship with the practitioners, to have their preferences respected, and to receive clear health information from the health care providers [56]. The use of VRIDAs within the PHV process was seen as having promising potential to improve patient/practitioner communication and collaboration within practice.

OTs perceived that the VRIDAs may also reduce anxiety and empower patients. Therefore, the use of VRIDA is likely to encourage therapists to consider new mechanisms to promote health literacy, which is a key enabling factor for patients to be empowered, take ownership, and be involved in the decisions that are made about their care. Health literacy is defined as the ability to "access, understand, evaluate, and communicate information as a way to promote, maintain, and improve health

XSL•FO
RenderX

in various settings over the life-course" [57]. Traditionally, health literacy tools have taken the form of information leaflets and/or delivery of information verbally. Some of the benefits of using information leaflets have been seen to be that patients and service users are able to refer back to written health information when required and to use the information at their own pace [58]. However, a relationship also exists between poor literacy skills, poor health, and poor health outcomes [59,60]. The use of more visually focused health communication tools, such as VRIDAs, are likely to provide the opportunity to overcome some of the communication imbalances that exist in current practice settings. Indeed, one recent study exploring the use of a virtual reality application to assess whether it could be used for persons with intellectual disabilities to achieve improved provision and communication of health-related information has achieved very promising results [61].

VRIDA applications such as SweetHome 3D were perceived by OTs as having the potential to address miscommunications that typically occur as part of the PHV process, as it gives patients immediate visual feedback on proposed home adaptations. Therapists appeared to believe that patients may prefer visual aids to facilitate understanding as opposed to more traditional methods of communication. To date only one study appears to exist within the research literature which explores the use of visual aids, in the form of photographs, within the process of occupational therapy home modifications or provision of assistive technology. Daniel et al [62] used photographs of the patient's home to evaluate the validity and feasibility of onsite home assessments with fallers. This technique could have significant benefits for persons with literacy issues, particularly as health literacy is more common among older adults [63]. In addition, evidence-based guidelines published by the Gerontological Society of America [64] suggest that visual aids can help address hearing-related communication issues. Moreover, visual aids can reduce the need for complex verbal information and reduce the cognitive effort required to understand information.

All the participants in this study were able to use the software, thus giving support to the notion that the majority of therapists can utilize technology and not just Generation X [65]. An Australian study [66] found that while therapists did use technology in practice, they used it primarily for client contact, professional development, and professional networking rather than for therapy provision. On the whole, participants were positive about the VRIDA software but two participants perceived the application as needing further refinement. In this study, most of the participants identified features that could enhance the software. Some participants viewed it as "quite sterile" and as being in need of further customization in order to achieve a more personalized look and feel, such as adding objects to the modelled environment that had some personal meaning and would make the environment feel more lived in. Indeed, evidence suggests that much of the personal home is tied up with the association of the self and identity [67]. Issues were also associated with exact measurement and whether this could hinder the development and use of the tool. Measurement is an interesting area of practice within occupational therapy, as to date little is known about how therapists measure for assistive technology and/or the instructions that are given in order to carry out measurement tasks [68].

Unlike many VR studies, this research was not tied to a specific clinical context, condition, or desired outcome such as learning surgical reconstruction [69], client interventions concerning public speaking [19], or interventions with stroke patients [70]. However, the wide variety of suggested uses for the different client groups, suggested by the participants, indicates the potential versatility and applicability of VRIDAs such as SweetHome 3D. One therapist did not perceive it as being useful for persons with cognitive issues, although one study has utilized the Engaging Platform for Art Development (ePAD) for persons with dementia in creative occupations with some success [71]. This study showed that a sample of people employed in occupational therapy possessed computer skills and that the software of choice was indeed usable, as also seems to be the case with the participants in this study. While this may seem like a small step, establishing these key points echoes the work of Laver et al [70] and the opinions of Verdonck and Ryan [15]. Much VR research focuses on the patient and their interactions with the software, without mention of whether or not clinical staff can operate it and feel confident doing so.

Limitations

A limitation of this research is that a follow-up interview was not carried out separately to the trial session itself; however, interviews were carried out at the end of the think-aloud sessions. This provided them with a chance to share any additional comments and reflect on the experience of using the software application. Qi [72] suggested that a follow-up interview may also allow the participants to validate the researchers' interpretation of their think-aloud utterances. This study relied on the recruitment of participants who were self-motivated to learn how to use the software, which may indicate that they felt comfortable in their own computer skills prior to participating in this research. In order to attract participants with a wide range of computer literacy, invitations could have explicitly stated that participants who have low as well as high computer skills are welcome to take part in the study. However, the insights gained from this study do represent views of OTs who have significant experience across a wide range of clinical settings and application domains. The number of participants that took part in this study may be considered to be too small to make generalizations about OT perceptions of the use of VRIDA in the PHV process more generally. However, in accordance with recent research findings in the usability testing research domain, the number of participants that took part in this study exceeds the suggested threshold number of five participants that are necessary to provide useful and effective feedback when using the think-aloud protocol for interactive prototype evaluation [40]. In relation to the TAM model, it is noted that it has been advocated that the Human Activity Assistive Technology (HAAT) model integrates the social model of disability, concepts from occupational therapy theory, and principles of assistive technology adoption and abandonment. Hence, if HAAT was used in addition to TAM, additional insights relating to these aspects may have been identified [73]. Nevertheless, TAM is a well-recognized model that has been used with significant success to identify barriers

XSL•FO
RenderX

to adoption of new technology within health care and a variety of other settings and provided an appropriate framework through which issues relating to the adoption of VRIDA were identified in a considered way.

Conclusions

This study has gained valuable insights into the value and utility of using VRIDA software applications such as SweetHome 3D within the occupational therapy setting and more specifically within the PHV process. OTs appeared to be positive about the utilization of VRIDAs within a range of clinical settings and that it would serve as a valuable collaborative tool that could empower patients and facilitate more effective patient/practitioner engagement. The study also revealed that VRIDAs have the potential to facilitate decision making and could serve as a valuable tool to demonstrate ideas and put them into a visual context that is personalized and intuitive for the patient. Furthermore, using VRIDA could better facilitate shared decision making and empower patients to play more of a role in the decisions that are made about their care. This is especially important given the complex emotions that can be tied to conditions leading to home modifications or the need for equipment. Furthermore, many studies look at the patient experience without noting the experience from the point of view of the clinician. It is often assumed that clinicians have/do not have the ability to learn to use new technology in practice. Without gathering and documenting the clinician's perspective, research is missing the valuable insights that clinicians can bring as a result of their range of clinical experience and that can be fed back into the development of technology that is tailored to the clinicians needs. This study has identified a number of issues that now can be addressed in order to ensure that the proposed VRIDA technology is suitably adapted and made to be fit for purpose, if it is to be introduced as a tool to facilitate more effective PHV interventions. Ultimately, new tools and strategies that enable improved patient/practitioner communication and collaboration must be identified and deployed, if significant levels of equipment abandonment seen as a result of PHV interventions are to be addressed and overcome. The use of VRIDA as a tool to facilitate improved communication and collaboration within this process has been perceived to be promising by practitioners.

Further research is needed to explore patient perceptions of VRIDA and to better understand the effectiveness of using such applications jointly and collaboratively with patients and practitioners. Further development work is also needed to incorporate the requirements suggested by practitioners as a result of this study and to identify patient specific requirements, which will ensure that both patients and practitioners are able to optimally benefit from using this application in practice.

Conflicts of Interest

None declared.

References

1. Darzi A. The National Archives. London; 2008 Jun 30. High quality care for all: NHS next stage review final report URL: http://webarchive.nationalarchives.gov.uk/20130107105354/http:/www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_085825 [accessed 2014-11-27] [WebCite Cache ID 6UP1haGPj]
2. Newman J, Kuhlmann E. Evidence-informed patient choice: Practical issues of involving patients in decisions about health care technologies. Journal of European Social Policy 2007;17(2):99-111.
3. Department of Health. Equity and Excellence: Liberating the NHS. UK: Stationery Office; Jul 12, 2010.
4. Money A, Barnett J, Kuljis J. Public claims about automatic external defibrillators: an online consumer opinions study. BMC Public Health 2011;11:332 [FREE Full text] [doi: 10.1186/1471-2458-11-332] [Medline: 21592349]
5. Money AG, Barnett J, Kuljis J, Lucas J. Patient perceptions of epinephrine auto-injectors: exploring barriers to use. Scand J Caring Sci 2013 Jun;27(2):335-344. [doi: 10.1111/j.1471-6712.2012.01045.x] [Medline: 22834703]
6. Bodenheimer T, Lorig K, Holman H, Grumbach K. Patient self-management of chronic disease in primary care. JAMA 2002 Nov 20;288(19):2469-2475. [Medline: 12435261]
7. Lord Ashley of Stoke. House of Commons. London, UK: The Stationery Office; 2010. Health and Social Care (Independent Living) Bill URL: http://www.publications.parliament.uk/pa/ld200910/ldbills/019/2010019.pdf [accessed 2014-11-27] [WebCite Cache ID 6UP30Ir8R]
8. Atwal A, McIntyre A, Craik C, Hunt J. Occupational Therapists' Perceptions of Predischarge Home Assessments with Older Adults in Acute Care. British Journal of Occupational Therapy 2008;71(2):52-58.
9. Lysack C, Neufeld S. Occupational therapist home evaluations: inequalities, but doing the best we can? Am J Occup Ther 2003;57(4):369-379. [Medline: 12911078]
10. Atwal A, McIntyre A, Craik C, Hunt J. Older adults and carers' perceptions of pre-discharge occupational therapy home visits in acute care. Age Ageing 2008 Jan;37(1):72-76 [FREE Full text] [doi: 10.1093/ageing/afm137] [Medline: 17971390]
11. Mann WC, Goodall S, Justiss MD, Tomita M. Dissatisfaction and nonuse of assistive devices among frail elders. Assist Technol 2002;14(2):130-139. [doi: 10.1080/10400435.2002.10132062] [Medline: 14651251]
12. Nygård L, Grahn U, Rudenhammar A, Hydling S. Reflecting on practice: are home visits prior to discharge worthwhile in geriatric inpatient care? Scand J Caring Sci 2004 Jun;18(2):193-203. [doi: 10.1111/j.1471-6712.2004.00270.x] [Medline: 15147483]

XSL•FO
RenderX

13. Crabtree A, Hemmings T, Rodden T, Chevherst K, Clarke K, Dewsbury G, et al. Designing with care: adapting cultural probes to inform design in sensitive settings. Brisbane, Australia: Ergonomics Society of Australia; 2003 Presented at: OzCHI; 26-26 November 2003; Brisbane, Australia p. 4-13.
14. Atwa A, Money AG, Spiliotopoulou G, Mcintyre A. Occupational therapists' perceptions about the clinical utility of the 3D interior design software. Disabil Rehabil Assist Technol 2013 Jul;8(4):348-355. [doi: 10.3109/17483107.2012.713437] [Medline: 22992195]
15. Verdonck M, Ryan S. Mainstream technology as an occupational therapy tool: Technophobe or technogeek? The British Journal of Occupational Therapy 2008;71(6):253-256.
16. Celinder D, Peoples H. Stroke patients' experiences with Wii Sports® during inpatient rehabilitation. Scand J Occup Ther 2012 Sep;19(5):457-463. [doi: 10.3109/11038128.2012.655307] [Medline: 22339207]
17. Katz N, Ring H, Naveh Y, Kizony R, Feintuch U, Weiss PL. Interactive virtual environment training for safe street crossing of right hemisphere stroke patients with unilateral spatial neglect. Disabil Rehabil 2005 Oct 30;27(20):1235-1243. [doi: 10.1080/09638280500076079] [Medline: 16298925]
18. Sutcliffe A. Multimedia and Virtual Reality: Designing Multisensory User Interfaces. London, UK: Psychology Press; 2003.
19. Wallach H, Safir M, Bar-Zvi M. Virtual reality cognitive behavior therapy for public speaking anxiety. Behavior Modification 2009;33(3):314-338. [doi: 10.1177/0145445511429999]
20. Bruce M, Regenbrecht H. A virtual reality claustrophobia therapy system – implementation test. 2009 Presented at: IEEE Virtual Reality Conference; 14-18 March 2009; Lafayette, LA, USA p. 179-182. [doi: 10.1109/VR.2009.4811020]
21. Chong A, Park A, Lee J. Web-based 3D system for home design. In: Stephanidis C, editor. Universal Access in Human-Computer Interaction. Berlin, Germany: Springer-Verlag; 2009:29-38.
22. Palmon O, Oxman R, Shahar M, Weiss P. Virtual environments as an aid to the design and evaluation of home and work settings for people with physical disabilities. In: Computer Aided Architectural Design Futures. Berlin, Germany: Springer-Verlag; 2005 Presented at: 11th International CAAD Futures Conference; June 20-22, 2005; Vienna, Austria.
23. Yoon S, Oh H, Cho J. Understanding furniture design choices using a 3D virtual showroom. Journal of Interior Design 2010;35(3):33-50. [doi: 10.1111/j.1939-1668.2010.01041.x]
24. Barlow J, Wright C, Sheasby J, Turner A, Hainsworth J. Self-management approaches for people with chronic conditions: a review. Patient Educ Couns 2002;48(2):177-187. [Medline: 12401421]
25. Chau P, Hu P. Examining a model of information technology acceptance by individual professionals: an exploratory study. Journal of Management Information Systems 2002;18(1):191-229.
26. Karsh BT, Escoto KH, Beasley JW, Holden RJ. Toward a theoretical approach to medical error reporting system research and design. Appl Ergon 2006 May;37(3):283-295 [FREE Full text] [doi: 10.1016/j.apergo.2005.07.003] [Medline: 16182233]
27. Audet AM, Doty MM, Peugh J, Shamasdin J, Zapert K, Schoenbaum S. Information technologies: when will they make it into physicians' black bags? MedGenMed 2004 Dec 6;6(4):2 [FREE Full text] [Medline: 15775829]
28. Schaper LK, Pervan GP. ICT and OTs: a model of information and communication technology acceptance and utilisation by occupational therapists. Int J Med Inform 2007 Jun;76 Suppl 1:S212-S221. [doi: 10.1016/j.ijmedinf.2006.05.028] [Medline: 16828335]
29. Lee Y, Kozar K, Larsen K. The technology acceptance model: past, present, and future. Communications of the Association for Information Systems 2003;12:752-780 [FREE Full text]
30. 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]
31. Bagozzi R. The legacy of the technology acceptance model and a proposal for a paradigm shift. Journal of the Association of Information Systems 2007;8:244-254 [FREE Full text]
32. Davis F. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 1989;13(3):319-340 [FREE Full text]
33. Yarbrough A, Smith T. Technology acceptance among physicians: a new take on TAM. Med Care Res Rev 2007 Dec;64(6):650-672. [doi: 10.1177/1077558707305942] [Medline: 17717378]
34. Kowitlawakul Y. The technology acceptance model: predicting nurses' intention to use telemedicine technology (eICU). Comput Inform Nurs 2011 Jul;29(7):411-418. [doi: 10.1097/NCN.0b013e3181f9dd4a] [Medline: 20975536]
35. Yi M, Jackson J, Park J, Probst J. Understanding information technology acceptance by individual professionals: toward an integrative view. Information & Management 2006 Apr;43(3):350-363. [doi: 10.1016/j.im.2005.08.006]
36. Van Schaik P, Bettany-Saltikov J, Warren J. Clinical acceptance of a low-cost portable system for postural assessment. Behaviour & Information Technology 2002 Jan;21(1):47-57. [doi: 10.1080/01449290110107236]
37. Chang H. Application of the extended Technology Acceptance Model to picture achiving and communication systems in dental hospitals. J Korean Soc Med Inform 2009;15(3):265-272. [doi: 10.4258/jksmi.2009.15.3.265]
38. Lishan X, Chiuan YC, Choolani M, Chuan CH. The perception and intention to adopt female-focused healthcare applications (FHA): a comparison between healthcare workers and non-healthcare workers. Int J Med Inform 2009 Apr;78(4):248-258. [doi: 10.1016/j.ijmedinf.2008.07.014] [Medline: 18783983]

XSL•FO
RenderX

39. Wu P. Opening the black boxes of TAM: towards a mixed methods approach. 2009 Presented at: 30th International Conference on Information Systems; December 18, 2009; Phoenix, Arizona, USA p. 1-10.
40. Borsci S, Macredie R, Barnett J, Martin J, Kuljis J, Young T. Reviewing and Extending the Five-User Assumption. ACM Trans. Comput.-Hum. Interact 2013 Nov 01;20(5):1-23. [doi: 10.1145/2506210]
41. Nielsen J, Landauer T. A mathematical model of the finding of usability problems. 1993 Presented at: INTERACT '93 CHI '93 Conference on Human Factors in Computing Systems; April 24-29, 1993; Amsterdam, The Netherlands p. 206-213.
42. Hanington B, Martin H. Universal Methods of Design. Beverley, MA, UAS: Rockport; 2012.
43. Unsworth CA. Using a head-mounted video camera to study clinical reasoning. Am J Occup Ther 2001;55(5):582-588. [Medline: 14601821]
44. Cooke L. Assessing concurrent think-aloud protocol as a usability test method: a technical communication approach. IEEE Trans Profess Commun 2010 Sep;53(3):202-215. [doi: 10.1109/TPC.2010.2052859]
45. King N. Using templates in the thematic analysis of text. In: Symon G, Cassell C, editors. Essential Guide to Qualitative Methods in Organizational Research. London: SAGE Publications; 2004:256-270.
46. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
47. Madill A, Jordan A, Shirley C. Objectivity and reliability in qualitative analysis: realist, contextualist and radical constructionist epistemologies. Br J Psychol 2000 Feb;91 (Pt 1):1-20. [Medline: 10717768]
48. Boyatzis R. Transforming qualitative information: Thematic analysis and code development. Thousand Oaks, CA, USA: Sage Publications; 1998.
49. Joffe H. Content and thematic analysis. In: Marks DF, Yardley L, editors. Research Methods for Clinical and Health Psychology. London, UK: Sage Publications; 2004:56-66.
50. Kuperman GJ. Health-information exchange: why are we doing it, and what are we doing? J Am Med Inform Assoc 2011;18(5):678-682 [FREE Full text] [doi: 10.1136/amiajnl-2010-000021] [Medline: 21676940]
51. Moats G. Discharge decision-making, enabling occupations, and client-centred practice. Can J Occup Ther 2007 Apr;74(2):91-101. [Medline: 17458368]
52. Atwal A, Spiliotopoulou G, Stradden J, Fellows V, Anako E, Robinson L, et al. Factors influencing occupational therapy home visit practice: a qualitative study. Scand J Occup Ther 2014 Jan;21(1):40-47. [doi: 10.3109/11038128.2013.821162] [Medline: 24102559]
53. Iwarsson S, Wahl HW, Nygren C, Oswald F, Sixsmith A, Sixsmith J, et al. Importance of the home environment for healthy aging: conceptual and methodological background of the European ENABLE-AGE Project. Gerontologist 2007 Feb;47(1):78-84. [Medline: 17327543]
54. Money AG, Fernando S, Lines L, Elliman AD. Developing and evaluating web-based assistive technologies for older adults. Gerontechnology 2009;8(3):165-177. [doi: 10.4017/gt.2009.08.03.013.00]
55. Money AG, Lines L, Fernando S, Elliman AD. e-Government online forms: design guidelines for older adults in Europe. Univ Access Inf Soc 2010 May 19;10(1):1-16. [doi: 10.1007/s10209-010-0191-y]
56. Bastiaens H, Van Royen P, Pavlic DR, Raposo V, Baker R. Older people's preferences for involvement in their own care: a qualitative study in primary health care in 11 European countries. Patient Educ Couns 2007 Sep;68(1):33-42. [doi: 10.1016/j.pec.2007.03.025] [Medline: 17544239]
57. Levasseur M, Carrier A. Integrating health literacy into occupational therapy: findings from a scoping review. Scand J Occup Ther 2012 Jul;19(4):305-314 [FREE Full text] [doi: 10.3109/11038128.2011.588724] [Medline: 21702741]
58. Johnson A, Sandford J. Written and verbal information versus verbal information only for patients being discharged from acute hospital settings to home: systematic review. Health Educ Res 2005 Aug;20(4):423-429 [FREE Full text] [doi: 10.1093/her/cyg141] [Medline: 15572437]
59. Dewalt DA, Berkman ND, Sheridan S, Lohr KN, Pignone MP. Literacy and health outcomes: a systematic review of the literature. J Gen Intern Med 2004 Dec;19(12):1228-1239 [FREE Full text] [doi: 10.1111/j.1525-1497.2004.40153.x] [Medline: 15610334]
60. Bostock S, Steptoe A. Association between low functional health literacy and mortality in older adults: longitudinal cohort study. BMJ 2012 Mar 15;344(mar15 3):e1602-e1602. [doi: 10.1136/bmj.e1602]
61. Hall V, Conboy-Hill S, Taylor D. Using virtual reality to provide health care information to people with intellectual disabilities: acceptability, usability, and potential utility. J Med Internet Res 2011;13(4):e91 [FREE Full text] [doi: 10.2196/jmir.1917] [Medline: 22082765]
62. Daniel H, Oesch P, Stuck A, Born S, Bachmann S, Schoenenberger A. Evaluation of a novel photography-based home assessment protocol for identification of environmental risk factors for falls in elderly persons. Swiss Med Wkly 2013;143:w13884 [FREE Full text] [doi: 10.4414/smw.2013.13884] [Medline: 24222638]
63. Chew LD, Bradley KA, Flum DR, Cornia PB, Koepsell TD. The impact of low health literacy on surgical practice. Am J Surg 2004 Sep;188(3):250-253. [doi: 10.1016/j.amjsurg.2004.04.005] [Medline: 15450829]
64. The Gerentological Society of America. Communicating with older adults an evidence-based review of what really works. Washington DC, USA; 2012. URL: http://www.agingresources.com/cms/wp-content/uploads/2012/10/GSA_Communicating-with-Older-Adults-low-Final.pdf [accessed 2014-11-25] [WebCite Cache ID 6ULfK8S8x]

XSL•FO
RenderX

65. Hills C, Ryan S, Warren-Forward H, Smith DR. Managing 'Generation Y' occupational therapists: optimising their potential. Aust Occup Ther J 2013 Aug;60(4):267-275. [doi: 10.1111/1440-1630.12043] [Medline: 23888977]
66. Chedid RJ, Dew A, Veitch C. Barriers to the use of Information and Communication Technology by occupational therapists working in a rural area of New South Wales, Australia. Aust Occup Ther J 2013 Jun;60(3):197-205. [doi: 10.1111/1440-1630.12016] [Medline: 23730785]
67. Gillsjö C, Schwartz-Barcott D, von Post I. Home: The place the older adult can not imagine living without. BMC Geriatr 2011;11(1):10. [doi: 10.1186/1471-2318-11-10]
68. Money A, Atwal A, McIntyre A, Elliman AD, French T. Bringing the home into the hospital: assisting the pre-discharge home visit process using 3D home visualization software. In: Stephanidis C, editor. Universal Access in Human-Computer Interaction. Applications Services. Heidelberg: Springer-Verlag; 2011:416-426.
69. Suero EM, Hüfner T, Stübig T, Krettek C, Citak M. Use of a virtual 3D software for planning of tibial plateau fracture reconstruction. Injury 2010 Jun;41(6):589-591. [doi: 10.1016/j.injury.2009.10.053] [Medline: 19939376]
70. Laver K, George S, Ratcliffe J, Crotty M. Virtual reality stroke rehabilitation--hype or hope? Aust Occup Ther J 2011 Jun;58(3):215-219. [doi: 10.1111/j.1440-1630.2010.00897.x] [Medline: 21599688]
71. Leuty V, Boger J, Young L, Hoey J, Mihailidis A. Engaging older adults with dementia in creative occupations using artificially intelligent assistive technology. Assist Technol 2013;25(2):72-79. [Medline: 23923689]
72. Qi DS. An Inquiry into Language-switching in Second Language Composing Processes. Canadian Modern Language Review/ La Revue canadienne des langues vivantes 2006 Sep 19;54(3):413-435. [doi: 10.3138/cmlr.54.3.413]
73. Giesbrecht E. Application of the Human Activity Assistive Technology model for occupational therapy research. Aust Occup Ther J 2013 Aug;60(4):230-240. [doi: 10.1111/1440-1630.12054] [Medline: 23888973]

Abbreviations

AU: actual use
HAAT: human activity assistive technology
ICT: information and communication technology
NHS: National Health Service
OT: occupational therapist
PDA: personal digital assistant
PEOU: perceived ease of use
PHV: pre-discharge home visit
PU: perceived usefulness
TAM: technology acceptance model
VR: virtual reality
VRIDA: virtual reality interior design application

Edited by G Eysenbach; submitted 25.07.14; peer-reviewed by I Paraskevopoulos, D Daylamani-Zad, N Plastow; comments to author 13.08.14; accepted 09.09.14; published 18.12.14.

Please cite as:
Atwal A, Money A, Harvey M
Occupational Therapists' Views on Using a Virtual Reality Interior Design Application Within the Pre-Discharge Home Visit Process
J Med Internet Res 2014;16(12):e283
URL: http://www.jmir.org/2014/12/e283/
doi:10.2196/jmir.3723
PMID:25526615

©Anita Atwal, Arthur Money, Michele Harvey. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 18.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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 Search Engine to Access PubMed Monolingual Subsets: Proof of Concept and Evaluation in French

Nicolas Griffon[1,2,3,4], MD, PhD; Matthieu Schuers[1,5], MD; Lina Fatima Soualmia[1,2], PhD; Julien Grosjean[1], MSc; Gaétan Kerdelhué[1]; Ivan Kergourlay[1], MSc; Badisse Dahamna[1], MSc; Stéfan Jacques Darmoni[1,2], MD, PhD

[1]CISMeF, TIBS, LITIS EA 4108, Rouen University Hospital, Normandy, Rouen, France

[2]U1142, LIMICS, Inserm, Paris, France

[3]UMR_S 1142, LIMICS, UPMC Univ Paris 06, Sorbonne Universités, Paris, France

[4]LIMICS (UMR_S 1142), Univ. Paris 13, Sorbonne Paris Cité, Villetaneuse, France

[5]Département de Médecine Générale, Université de Rouen, Rouen, France

Corresponding Author:
Nicolas Griffon, MD, PhD
CISMeF
TIBS, LITIS EA 4108
Rouen University Hospital, Normandy
Unité d'informatique clinique - batiment Pillore
1 rue de Germont
Rouen, 76031
France
Phone: 33 232885726
Fax: 33 232888909
Email: nicolas.griffon@chu-rouen.fr

Abstract

Background: PubMed contains numerous articles in languages other than English. However, existing solutions to access these articles in the language in which they were written remain unconvincing.

Objective: The aim of this study was to propose a practical search engine, called Multilingual PubMed, which will permit access to a PubMed subset in 1 language and to evaluate the precision and coverage for the French version (Multilingual PubMed-French).

Methods: To create this tool, translations of MeSH were enriched (eg, adding synonyms and translations in French) and integrated into a terminology portal. PubMed subsets in several European languages were also added to our database using a dedicated parser. The response time for the generic semantic search engine was evaluated for simple queries. BabelMeSH, Multilingual PubMed-French, and 3 different PubMed strategies were compared by searching for literature in French. Precision and coverage were measured for 20 randomly selected queries. The results were evaluated as relevant to title and abstract, the evaluator being blind to search strategy.

Results: More than 650,000 PubMed citations in French were integrated into the Multilingual PubMed-French information system. The response times were all below the threshold defined for usability (2 seconds). Two search strategies (Multilingual PubMed-French and 1 PubMed strategy) showed high precision (0.93 and 0.97, respectively), but coverage was 4 times higher for Multilingual PubMed-French.

Conclusions: It is now possible to freely access biomedical literature using a practical search tool in French. This tool will be of particular interest for health professionals and other end users who do not read or query sufficiently in English. The information system is theoretically well suited to expand the approach to other European languages, such as German, Spanish, Norwegian, and Portuguese.

(J Med Internet Res 2014;16(12):e271) doi:10.2196/jmir.3836

KEYWORDS

databases, bibliographic; French language; information storage and retrieval; PubMed; user-computer interface; search engine

XSL•FO
RenderX

Introduction

MEDLINE, created by the US National Library of Medicine (NLM), is the most used medical bibliographic database in the world. Currently (as of September 3, 2014), it contains 21,515,657 citations [1] from 5650 indexed journals from 81 countries around the world. Each MEDLINE record is indexed with the NLM's controlled vocabulary, Medical Subject Headings (MeSH) [2].

MEDLINE is the largest component of PubMed [3], the freely accessible online database of biomedical journal citations and abstracts. In addition to MEDLINE citations, PubMed also contains [4]:

1. Citations not yet indexed with MeSH and added to MEDLINE (in-process citations or when supplied electronically by the publisher);
2. Some "old MEDLINE" citations that have not yet been updated with the current vocabulary and converted to MEDLINE status;
3. Citations to some additional life science journals that submit full text to PubMedCentral and receive a qualitative review by the NLM; and
4. Citations to author manuscripts of articles published by NIH-funded researchers.

On the same date (September 3, 2014), PubMed contained 24,157,837 citations [5].

Language may be an obstacle to access PubMed and all the information it contains [6]. Moreover, there is still demand for native-language articles [7]. Of the 28 different ways to access PubMed listed by Lu [8], only 2 help non-native-English speakers to query PubMed/MEDLINE in their native language: BabelMeSH [9,10] and Patient, Intervention, Comparison, Outcome (PICO) Linguist [10]. We developed a French MeSH browser linked to PubMed [11]; more than 500 users consult it daily and it is taught in half of French medical schools. These 3 tools rely on MeSH translation in multiple languages to ease querying [12] and also rely on some metadata available in existing user languages to ease the results of browsing. Nevertheless, these 3 tools lack most PubMed functionalities (eg, advanced query builder, filters, citation sensor); in truth, they just lack functionalities.

A multilingual search engine to access the PubMed/MEDLINE subset in any non-English language (eg, French, German, Spanish, or Norwegian) with advanced functionalities would be of great interest for any user who is not comfortable in English. The goal of this paper was to present such a tool (Multilingual PubMed) and evaluate the potential increase in information retrieval task performance by using the French version.

Methods

Material

To create an effective multilingual search engine, the basic framework must first be clearly specified. Several tools were initially developed before the beginning of this project. Since 1995, the Catalogue and Index of Online Health Resources in French (CISMeF=Catalogue et Index des Sites Médicaux de langue Française), a quality-controlled health gateway [13], describes and indexes using the MeSH thesaurus the main institutional health Web resources (documents and sites) available in French. A search engine, Doc'CISMeF, was developed in 2000 to allow querying of this gateway. In 2007, CISMeF began using several terminologies/ontologies to allow easier and more accurate indexing. The CISMeF team was prompted to develop a multiterminology portal (French/English) [14], which interoperated with the search engine. In 2012, a transition was made to establish a cross-lingual portal named Health Terminology/Ontology Portal (HeTOP) available in 23 different languages [15,16].

Recently, Doc'CISMeF was significantly updated in 2 ways: (1) it has become a generic tool able to integrate any metadata model and its related data; therefore, this tool is able to describe and index Web resources as well as PubMed citations and (2) the search engine has now been developed into a reliable multilingual tool, which is capable of searching by using multiple terminologies/ontologies in several languages.

These 2 improvements in Doc'CISMeF allow the search for PubMed citation in any language. The final steps to transform Doc'CISMeF into the desired Multilingual Access to Each PubMed Subset for Each Language (Multilingual PubMed or, for French, Multilingual PubMed-French), designed for non-English speakers to access the literature in languages other than English, consist of (1) completing translation of MeSH descriptors and MeSH qualifiers and (2) translating the interface.

For the 5 basic languages of this study, the translation of MeSH descriptors is available in Unified Medical Language System [17]. Several institutions have helped us to translate the remaining MeSH qualifiers and interface terms (see Acknowledgments).

Nevertheless, developing a search engine requires good knowledge of the language used. As a French team, we are unable to assume that the performance for all other languages is equivalent to that of French language.

Two improvements related to the MeSH thesaurus in Doc'CISMeF can be used in Multilingual PubMed [18]:

1. Bilingual (French and English) predefined queries (n=389) were defined by the CISMeF medical librarians (eg, the CISMeF term "natremia" was automatically transformed to the MeSH predefined query "sodium/blood"). These predefined queries will have to be translated to be applied to the other European languages of this project.
2. Superconcepts (or metaterms; n=126) are medical (sub)specialties or biological science disciplines (eg, cardiology or bacteriology) selected by the CISMeF chief librarian. For each superconcept, 1 semantic link was created with 1 or more MeSH descriptor and qualifier. For example, the metaterm *psychiatry* is associated with the MeSH descriptors *psychiatry* and *psychiatric hospital* that belong to a completely different tree structure within the MeSH. Superconcepts have been created to optimize information retrieval and to overcome the relatively

XSL•FO
RenderX

restrictive nature of MeSH descriptors. These superconcepts also need to be translated, but the basic semantic links between superconcepts and MeSH descriptors are language-independent.

The list of predefined queries and metaterms is available at the HeTOP cross-lingual terminologies/ontologies portal [15].

There are also some other improvements that are based on all the available metadata. It is possible to perform faceted browsing [19]. Facets list citation characteristics, which can be used to refine results. It works based on metadata (eg, year, country, resource type). Default ranking in Doc'CISMeF is based on a relevance score, whereas in PubMed, the most recent articles are ranked first: last in-first out (LIFO). Doc'CISMeF interprets users' queries to identify meaningful words and medical descriptors matching health terminologies, such as MeSH. It then returns results corresponding to these words and descriptors. It uses several criteria to score and rank the results: query words present in the title, descriptors present in major topics, publication date, and origin of the descriptors as they may have been assigned by librarians or programmatically by computers for less important resources. If title words and major topics perfectly matched the user's query, they score 100% and are consequently ranked. Imperfect matches score less depending on how many words and descriptors are present (eg, MeSH minor topic). For results of equal scores, publication date is the determinant criterion.

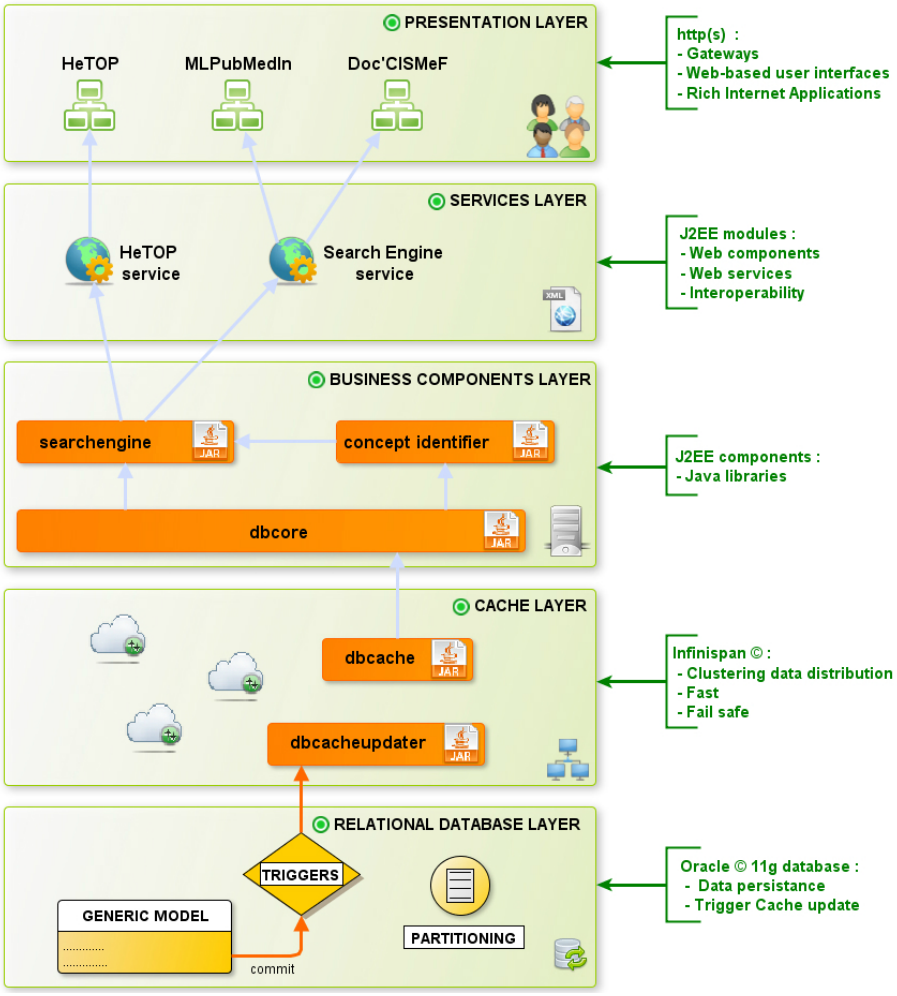
Some other improvements previously developed in Doc'CISMeF have not yet been implemented in the Multilingual PubMed-French search engine (see Discussion).

Architecture

The CISMeF generic information system is divided into 5 layers (Figure 1):

1. A relational database layer, which ensures data persistence and Structured Query Language query processing; execution plans are optimized by dynamic partitioning and local indexes.
2. A cache layer, which is devoted to clustering data distribution. It is based on random access memory and is distributed between several servers; therefore, it is very fast. Moreover, it avoids data loss after crashes.
3. A business components layer mainly composed of CISMeF Java libraries dealing with business logic.
4. A service layer hosting Web services that rely on business components and are, in turn, used by each client application of the information system.
5. A presentation layer that includes the gateways, Web-based user interfaces, rich Internet apps, and any client application of the service layer.

Figure 1. CISMeF information system.



XSL•FO
RenderX

Data Sources

The first and only version of Multilingual PubMed tested in this paper was the French version: the Multilingual PubMed-French search engine. Nevertheless, the whole framework should be able to manage any language as it does with French. To perform this study, all PubMed citations in French were extracted from PubMed and included in the Multilingual PubMed-French semantic search engine using a specific parser. Some PubMed citations were also extracted in other languages (ie, German, Spanish, Portuguese, and Norwegian) to ensure that our hypothesis of a multilingual framework was not unrealistic.

All MeSH descriptors and MeSH qualifiers were translated into French by the Technical and Scientific Information Department of the French National Institute of Health [20]. The CISMeF team considerably enriched this translation with:

1. 25,501 synonyms and 689 ambiguous acronyms of MeSH descriptors;
2. 163 synonyms of MeSH qualifiers;
3. 20,887 translations of MeSH supplementary concepts (out of 209,326);
4. 27,295 synonyms of MeSH supplementary concepts;
5. 6037 translations of MeSH scope notes, which consist of MeSH descriptor definitions; and
6. 3918 “see also” relations among MeSH descriptors.

Evaluation

The 2 main limiting factors for such a tool are the response time, which must be low in order to be acceptable for end users, and the result quality (are there any results and are they relevant?). They were evaluated by measuring the response time and relevance of the first 20 results for 20 queries.

Queries were selected from the queries frequently run on Doc’CISMeF (log-analysis). Twenty queries were randomly selected from those that retrieved results, were run more than 50 times in the previous 100 days, and did not involve advanced syntax.

Each query was run on Multilingual PubMed-French, BabelMeSH, and PubMed; the exact sentence from our log was pasted in each search engine. The corpus was limited to articles only in French language for each search engine. For PubMed, a second query was launched on the transliterated title (TT) field, hereafter referred to as the PubMed TT strategy. Finally, 2 PubMed algorithms for result sorting were tested: the classic LIFO and the recently proposed relevance ranking [21], which is probably more comparable to the Multilingual PubMed-French sorting algorithm. Thus, 5 strategies were tested: BabelMeSH, Multilingual PubMed-French, PubMed LIFO, PubMed relevance, and PubMed TT. Examples of queries are available in Table 1. PICO linguist was not evaluated because it works basically like Babel MeSH, but with a PICO interface (ie, with structured fields for patient, intervention, comparison, and outcome) which is designed for a specific query. The PubMed TT strategy was only studied with LIFO ranking because relevance ranking is based on the field in which the terms are found [21] and the PubMed TT strategy involved only 1 field: transliterated title. According to our data, PubMed LIFO and PubMed relevance ranking share 45% of their results. For PubMed TT strategy, this proportion rose to 88%.

The PubMed identifier (PMID) of the first 20 citations and the total number of citations were recorded by NG. For Multilingual PubMed-French, only response times were recorded. A medical librarian (GK) performed an assessment of citation relevance using a 3-modality Likert scale. He was blind to the search strategy that retrieved the resources to be evaluated. Therefore, it is possible to estimate the precision of each tool (ie, the proportion of relevant results in the results).

The relation between response time and number of citations was studied using the Pearson correlation coefficient. Two analyses were performed: for the strict one, only citations assessed as totally relevant were considered as true positives; for the relaxed one, both totally relevant and partially relevant citations were considered as true positives.

Table 1. Examples of queries in the different tools.

Search strategy	Example of query (for “constipation”)[a]
BabelMeSH[b]	constipation
Multilingual PubMed-French	constipation
PubMed[c]	constipation[All Fields] AND French[lang]
PubMed TT	constipation[TT] AND French[lang]

[a] Shown is the query as written in the search box.

[b] The checkbox “français” was checked.

[c] Whatever the ranking used (relevance or LIFO), the query was the same.

Results

Technical Feasibility

Table 2 shows the number of PubMed citations for the main European languages (excluding English). As a proof of concept, PubMed citations of different languages were included in the CISMeF information system: French (n=665,359), German (n=7102), Portuguese (n=4497), Spanish (n=4297), and Norwegian (n=3764). The end user was able to choose 1 language and then perform a query in the same language. Thanks to our colleagues mentioned in the Acknowledgments, the Multilingual PubMed interface was translated into French, German, Portuguese, Spanish, and Norwegian. The same main metadata displayed by default in the PubMed database was similarly displayed in the Multilingual PubMed search engine,

XSL•FO
RenderX

as well as the indexing MeSH descriptors and qualifiers (Figure 2). Direct links to the full-text article in the same language were also presented to the end user (Figure 2), often via Digital Object Identifier [22]. The objective of creating a bibliographic database extracted from PubMed in several languages and available in 1 specific language was then completed as a proof of concept.

In the previous CISMeF gateway, the objective for response time was the following: all queries with 1 or 2 terms should be less than 2 seconds. Since 2001, this objective has been fulfilled for the 101,000 Web resources included in the CISMeF information system. The number of PubMed citations in French versus CISMeF Web resources is approximately 6 times more. All response times for all queries were below the limit of 2 seconds (Table 3). As expected, there was a strong correlation between the response time and the number of citations retrieved (r=.73, 95% CI .42-.89). According to linear regression coefficient, this would lead to an unacceptable response time for queries that collect more than 20,000 citations. This evaluation of response time is an important step in determining the feasibility of this Multilingual PubMed-French search engine to create a bibliographic database in French for health professionals. The current version of Multilingual PubMed-French is available on the Internet [23].

Table 2. Number of PubMed citations for the main European languages.

Language	Number of PubMed citations
German	808,653
French	680,451[a]
Italian	294,720
Spanish	302,287
Portuguese	85,839
Norwegian	35,712

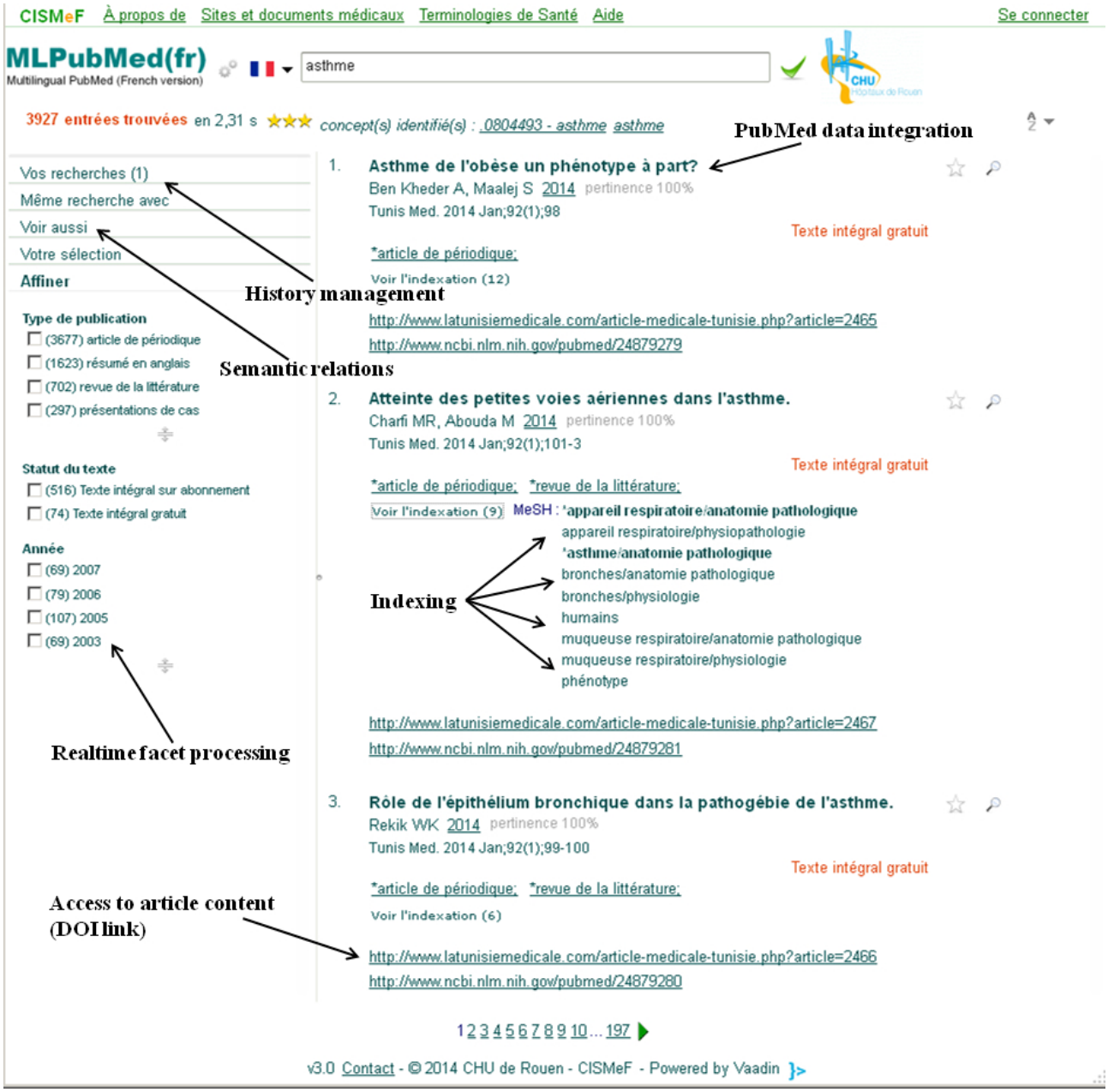
[a] The discrepancy between the number of French citations in PubMed and the number of French citations inserted in our information system comes primarily from some doubtful classifications in PubMed (ie, articles mislabeled as French).

Table 3. Response times for the 20 Multilingual PubMed-French queries.

Query (French/English)	Response time (seconds)	Number of PubMed citations
Allaitement maternel / breast feeding	0.42	633
Angine/pharyngitis	0.53	845
AVC/stroke	0.54	1704
BPCO/COPD	0.49	1503
Cigarette électronique / electronic cigarette	0.42	24
Constipation/constipation	0.27	531
Coqueluche / whooping cough	0.54	314
Gale/scabies	0.57	148
Hémochromatose/hemochromatosis	0.59	712
Hypertension/hypertension	1.03	8694
Lupus/lupus	0.34	1147
Maladie cœliaque / celiac disease	0.58	723
Maladie de Crohn / Crohn disease	0.69	1363
Nutrition/nutrition	1.13	8367
Psychiatrie/psychiatry	0.94	5602
Sarcoïdose/sarcoidosis	0.77	1721
Scoliose/scoliosis	1.10	669
Soins infirmiers / nursing care	0.86	5767
Tabac/tobacco	0.42	679
Toxoplasmose/toxoplasmosis	0.74	1238

XSL•FO
RenderX

Figure 2. Screenshot of the multilingual search engine.



Performance Evaluation

Table 4 displays the coverage of each query for each strategy. The PubMed TT strategy retrieved far fewer citations than the other 4 strategies with a total of 10,716 hits. Babel MeSH, Multilingual PubMed-French, and regular PubMed queries retrieved 50,894, 42,384, and 34,047 citations, respectively. Babel MeSH and Multilingual PubMed-French retrieved significantly more citations than relevance/LIFO PubMed (both P=.03; Mann-Whitney test).

The precision measured on the first 20 citations retrieved are listed (1) considering the citations assessed as highly relevant—the strict analysis—and (2) considering both the highly and the slightly relevant citations, the relaxed analysis (Table 5). The results were similar with 2 strategies that outperformed the other 3 in terms of precision: Multilingual PubMed-French and PubMed TT. These 2 strategies reached more than 90% precision with a slight advantage for PubMed TT (P=.002, Fisher test), whereas the other 3 barely reached 80% in the relaxed analysis.

XSL•FO
RenderX

Table 4. Coverage of queries according to the search engine.

Query (English)	Strategy (n)			
	BabelMeSH	Multilingual PubMed-French	PubMed[a]	PubMed TT
Breast feeding	639	633	3	50
Pharyngitis	2043	845	2326	82
Stroke	2671	1704	202	52
COPD	687	1503	25	275
Electronic cigarette	355	24	0	0
Constipation	574	531	666	144
Whooping cough	227	314	2	166
Scabies	136	148	4	102
Hemochromatosis	537	712	756	129
Hypertension	10,348	8694	12,580	1533
Lupus	2232	1147	2694	1150
Celiac disease	649	723	2	305
Crohn disease	1190	1363	2	758
Nutrition	4969	8367	6480	819
Psychiatry	4453	5602	5394	1940
Sarcoidosis	1522	1721	13	878
Scoliosis	651	669	3	159
Nursing care	14,360	5767	2867	888
Tobacco	1607	679	18	585
Toxoplasmosis	1044	1238	10	701
Total	50,894	42,384	34,047	10,716

[a] Whatever the ranking used (relevance or LIFO), the number of citations retrieved was the same.

XSL•FO
RenderX

Table 5. Precision measured on the first 20 citations for each query for each strategy (analysis-relaxed analysis).

Query (English)	Strategy (precision)				
	Babel MeSH	Multilingual PubMed-French	PubMed LIFO	PubMed Relevance	PubMed TT
Breast feeding	0.55-0.75	1.00-1.00	1.00-1.00[a]	1.00-1.00[a]	1.00-1.00
Pharyngitis	0.00-0.00	0.65-0.70	0.00-0.00	0.00-0.00	0.50-0.60
Stroke	0.45-0.80	0.95-0.95	0.15-0.15	0.35-0.35	1.00-1.00
COPD	0.40-0.55	1.00-1.00	0.75-0.85	1.00-1.00	1.00-1.00
Electronic cigarette	0.10-0.10	0.35-0.40	—[b]	—[b]	—[b]
Constipation	0.45-1.00	1.00-1.00	0.50-0.95	0.75-1.00	1.00-1.00
Whooping cough	0.80-1.00	1.00-1.00	1.00-1.00[c]	1.00-1.00[c]	1.00-1.00
Scabies	0.90-1.00	1.00-1.00	0.25-0.25[d]	0.25-0.25[d]	1.00-1.00
Hemochromatosis	0.65-0.95	1.00-1.00	0.70-1.00	0.85-1.00	1.00-1.00
Hypertension	0.45-0.75	0.95-0.95	0.25-0.75	1.00-1.00	1.00-1.00
Lupus	0.70-1.00	1.00-1.00	0.75-0.95	1.00-1.00	1.00-1.00
Celiac disease	0.70-0.95	1.00-1.00	1.00-1.00[c]	1.00-1.00[c]	1.00-1.00
Crohn disease	0.80-0.90	1.00-1.00	0.50-1.00[c]	0.50-1.00[c]	1.00-1.00
Nutrition	0.45-0.70	0.70-1.00	0.25-0.60	1.00-1.00	1.00-1.00
Psychiatry	0.75-0.80	1.00-1.00	0.85-0.95	0.90-0.95	1.00-1.00
Sarcoidosis	0.70-0.90	1.00-1.00	0.92-0.92[e]	0.92-0.92[e]	1.00-1.00
Scoliosis	0.65-0.90	1.00-1.00	0.67-0.67[a]	0.67-0.67[a]	1.00-1.00
Nursing care	0.85-0.90	1.00-1.00	0.70-0.80	0.85-0.90	1.00-1.00
Tobacco	0.60-0.90	1.00-1.00	0.83-0.89[f]	0.83-0.89[f]	1.00-1.00
Toxoplasmosis	0.65-0.90	1.00-1.00	1.00-1.00[g]	1.00-1.00[g]	1.00-1.00
Total	0.58-0.79	0.93-0.95	0.57-0.74	0.79-0.83	0.97-0.98

[a] Only 3 citations retrieved.
[b] No citation retrieved.
[c] Only 2 citations retrieved.
[d] Only 4 citations retrieved.
[e] Only 13 citations retrieved.
[f] Only 18 citations retrieved.
[g] Only 10 citations retrieved.

Discussion

Principal Results

The Multilingual PubMed-French strategy allows the retrieval of French PubMed citations with high precision and high coverage. Despite more than 600,000 PubMed citations that were introduced into the CISMeF information system, the average server-side response time stayed below the threshold time (ie, 2 seconds), which is acceptable, in our experience, for the average health professional. The Multilingual PubMed-French search engine is now freely available on the Internet [23].

Babel MeSH allows users to access PubMed database in French. It provides many results for low precision and is, therefore, of poor interest for French speakers. Moreover, the functioning of the search engine is not detailed anywhere; it is only possible to make rough free-term queries that are translated automatically by Babel MeSH. The Babel MeSH interface is limited; it is not possible to perform many queries in a short period of time.

The classic use of PubMed is not satisfying for French speakers, which is not surprising because it is supposed to work in English, not in French. The relatively high mean coverage masked enormous heterogeneity (from 0 to more than 12,000 citations). Precision was lower than that for Multilingual PubMed-French or PubMed TT. However, the relevance ranking provided consistently better precision for the first 20 results than LIFO ranking. There was no evidence to maintain LIFO ranking as default.

XSL•FO
RenderX

Using PubMed TT provides more precise results than using Multilingual PubMed-French. PubMed TT works like a free-text search engine that looks only in the title of the citation. Therefore, it is quite logical that precision was high. Nevertheless, PubMed TT recall must be lower than Multilingual PubMed-French recall based on the huge differences in coverage and slight differences in precision. Moreover, Multilingual PubMed-French and Babel MeSH constantly provided more than 20 results, whereas PubMed did not; some queries provided few or even no results. The poor coverage of PubMed TT is probably the consequence of difficulties managing French particularities (eg, elision, apostrophe, accentuation).

Limitations

This study has several limitations. First, this feasibility study only included server-side response time. The overall response time also depended on the end user's computer type and type of browser. Second, the MeSH thesaurus is incompletely translated. MeSH descriptors and qualifiers are frequently translated, but this is not the case for MeSH supplementary concepts. Therefore, querying at a MeSH supplementary concept level of accuracy is currently not possible in one's own language. It would require a huge amount of work to translate all MeSH supplementary concepts, even using natural language processing tools [24]. Last, English is by far the most frequently used language of scientific medical publications and publications in other languages are declining comparatively. Nevertheless, the latter are important because they can be more easily adapted to national contexts and be more understandable to a wider local audience. An automatic intelligible translation of English articles would definitely resolve this issue, but it is far beyond current technology [25].

For the "angine/pharyngitis" query, an unanticipated issue arose: PubMed (LIFO and relevance strategies) did not find any results and automatically translated the query to "angina" which has a different meaning. This affected the precision of these strategies by biasing it toward zero. Three sensitivity analyses were performed to evaluate bias importance: (1) the "angine" query was excluded from the analysis, (2) the "angine" query was considered as providing zero results for PubMed LIFO and PubMed relevance strategies, and (3) the 20 citations returned by this query for PubMed LIFO/relevance strategies were considered as fully relevant. In each scenario, the precision was significantly better for Multilingual PubMed-French and PubMed TT compared with PubMed LIFO/relevance strategies. Therefore, it is possible to conclude a very limited bias.

Perspectives and Future Challenges

One of the strengths of Multilingual PubMed lies in the multilingual CISMeF information system. Because the MeSH thesaurus is already translated into multiple languages, it will be easy to make Multilingual PubMed for these languages. Nevertheless, the CISMeF team is not familiar with the linguistic treatments that are necessary to interpret user queries in other languages than French; therefore, it is unlikely that the other versions of Multilingual PubMed will achieve results comparable to those observed here. We are currently working with several international teams to extend Multilingual PubMed to German, Spanish, Portuguese, and Norwegian. According to the number of PubMed citations (see Table 1) for the main European languages, this feasibility study could be extrapolated as positive for other languages with fewer PubMed citations than French (ie, Spanish, Italian, Portuguese, and Norwegian) and one-third larger for German.

Several other improvements related to MeSH indexing in the search engine have not yet been implemented for PubMed citations; in particular, affiliation of MeSH qualifiers to MeSH supplementary concepts, indexing with MeSH concepts, and affiliation of MeSH qualifiers to MeSH concepts [26]. It will not be feasible to perform these tasks manually due to the size of the PubMed citations corpus in French (n=665,359). Therefore, the automatic indexing tool used in the CISMeF quality-controlled health gateway will need to be enhanced to be applied to the Multilingual PubMed-French search engine.

As stated in the Introduction, it is possible to manage multiple kinds of resources in the CISMeF information system. End users are now able to ask for mixed results: Web resources from CISMeF and PubMed citations from Multilingual PubMed. Also, an interesting avenue for future development of Multilingual PubMed is to integrate bibliographic data from other sources. A partnership with Elsevier-Masson and 2 small French companies has now been launched. Elsevier will provide additional metadata for MEDLINE/PubMed citations (abstracts in French in particular), which currently are not yet available in the MEDLINE/PubMed database. Furthermore, Elsevier-Masson will provide metadata on journals not included in PubMed, but included in other bibliographic/bibliometric databases (eg, Web of Science, EMBASE, and BIOSIS).

In addition to being more efficient than its competitors are, Multilingual PubMed is, in our opinion, more user friendly and provides more features. Nevertheless, there are still many interesting functionalities in PubMed (ie, related citations, history management) or in other tools [8] (graphical representation) that are not implemented in Multilingual PubMed. Future work may address these issues.

Conclusion

This feasibility study was carried out to create a multilingual search engine to query monolingual PubMed subsets. It was successful for the French language and will be extended to the other main European languages. Bibliographical references from non-English publications can now be searched in a native language–friendly interface using multilingual data from the PubMed database and MeSH vocabulary.

Overall, this Multilingual PubMed tool could be of valuable interest for non–English-speaking health professionals who are unable to access PubMed.

XSL•FO
RenderX

Acknowledgments

The authors would like to thank Pablo Iriarte, Librarian at the University Hospital CHUV in Lausanne, Switzerland, for his decisive impulse for starting the project. The authors are grateful to Nikki Sabourin-Gibbs, Rouen University Hospital, for reviewing the manuscript in English. UMIT, Austria, Centre Tudor, Luxembourg, Hospital Italiano of Buenos Aires, Argentina, and the Norwegian Electronic Health Library graciously translated some MeSH terms and some interface terms into German, German and Portuguese, Spanish and Italian, and Norwegian, respectively.

Authors' Contributions

GK, LFS, and SJD had the idea of Multilingual PubMed and pushed for its development. JG, IK, and BD developed the whole application and optimized the architecture of the system. NG, GK, MS, and SJD conceived the evaluation. NG collected the data and performed analysis. GK made citation relevance assessment. NG and SJD drafted the manuscript. LFS and MS made substantial enhancement to it. All the authors approved the final manuscript.

Conflicts of Interest

None declared.

References

1. MEDLINE. Number of resources in MEDLINE URL: http://www.ncbi.nlm.nih.gov/pubmed?cmd=PureSearch&db=pubmed&term=medline[sb] [WebCite Cache ID 6S7HdKcK3]
2. Lipscomb CE. Medical Subject Headings (MeSH). Bull Med Libr Assoc 2000 Jul;88(3):265-266 [FREE Full text] [Medline: 10928714]
3. PubMed. URL: http://www.ncbi.nlm.nih.gov/pubmed/ [accessed 2014-08-26] [WebCite Cache ID 6S7LTaVwv]
4. US National Library of Medicine. 2014. Fact sheet: MEDLINE, PubMed, and PMC (PubMed Central): How are they different? URL: http://www.nlm.nih.gov/pubs/factsheets/dif_med_pub.html [accessed 2014-08-26] [WebCite Cache ID 6S7Moo8Q8]
5. PubMed. Number of resources in PubMed URL: http://www.ncbi.nlm.nih.gov/sites/entrez?cmd=PureSearch&db=pubmed&term=all[sb] [WebCite Cache ID 6S7Howfws]
6. Sheets L, Gavino A, Callaghan F, Fontelo P. Do language fluency and other socioeconomic factors influence the use of PubMed and MedlinePlus? Appl Clin Inform 2013;4(2):170-184 [FREE Full text] [doi: 10.4338/ACI-2013-01-RA-0006] [Medline: 23874356]
7. Chneiweiss H. [Why a scientific review in French?]. Med Sci (Paris) 2014 Jan;30(1):7-8. [doi: 10.1051/medsci/20143001001] [Medline: 24472447]
8. Lu Z. PubMed and beyond: a survey of web tools for searching biomedical literature. Database (Oxford) 2011;2011:baq036 [FREE Full text] [doi: 10.1093/database/baq036] [Medline: 21245076]
9. Liu F, Ackerman M, Fontelo P. BabelMeSH: development of a cross-language tool for MEDLINE/PubMed. AMIA Annu Symp Proc 2006:1012 [FREE Full text] [Medline: 17238631]
10. Fontelo P, Liu F, Leon S, Anne A, Ackerman M. PICO Linguist and BabelMeSH: development and partial evaluation of evidence-based multilanguage search tools for MEDLINE/PubMed. Stud Health Technol Inform 2007;129(Pt 1):817-821. [Medline: 17911830]
11. Thirion B, Pereira S, Névéol A, Dahamna B, Darmoni S. French MeSH Browser: a cross-language tool to access MEDLINE/PubMed. AMIA Annu Symp Proc 2007:1132. [Medline: 18694229]
12. Nelson SJ, Schopen M, Savage AG, Schulman JL, Arluk N. The MeSH translation maintenance system: structure, interface design, and implementation. Stud Health Technol Inform 2004;107(Pt 1):67-69. [Medline: 15360776]
13. Darmoni SJ, Leroy JP, Baudic F, Douyère M, Piot J, Thirion B. CISMeF: a structured health resource guide. Methods Inf Med 2000 Mar;39(1):30-35. [Medline: 10786067]
14. Portail Terminologique de Santé. URL: http://pts.chu-rouen.fr/old/ [accessed 2014-08-26] [WebCite Cache ID 6S7GWqPTy]
15. Grosjean J, Merabti T, Griffon N, Dahamna B, Darmoni SJ. Teaching medicine with a terminology/ontology portal. Stud Health Technol Inform 2012;180:949-953. [Medline: 22874333]
16. Health Terminology/Ontology Portal. URL: http://www.hetop.fr/hetop/ [accessed 2014-08-26] [WebCite Cache ID 6S7GyugRz]
17. Humphreys BL, Lindberg DA, Schoolman HM, Barnett GO. The Unified Medical Language System: an informatics research collaboration. J Am Med Inform Assoc 1998;5(1):1-11 [FREE Full text] [Medline: 9452981]
18. Douyère M, Soualmia LF, Névéol A, Rogozan A, Dahamna B, Leroy JP, et al. Enhancing the MeSH thesaurus to retrieve French online health resources in a quality-controlled gateway. Health Info Libr J 2004 Dec;21(4):253-261. [doi: 10.1111/j.1471-1842.2004.00526.x] [Medline: 15606883]
19. Tunkelang D. Faceted search. Synthesis Lectures on Information Concepts, Retrieval, and Services 2009 Jan;1(1):1-80. [doi: 10.2200/S00190ED1V01Y200904ICR005]

XSL•FO
RenderX

20. Le MeSH bilingue anglais - français. URL: http://mesh.inserm.fr/mesh/ [accessed 2014-08-26] [WebCite Cache ID 6S7MtSGFa]
21. Canese K. PubMed relevance sort. NLM Tech Bull 2013 Oct;394:e2.
22. Digital Object Identifier System. URL: http://www.doi.org/ [accessed 2014-08-26] [WebCite Cache ID 6S7MzU2OQ]
23. MLPubMedFr. URL: http://doccismef.chu-rouen.fr/dc/?env=pubmed [accessed 2014-08-26] [WebCite Cache ID 6S7HHSAyT]
24. Merabti T, Soualmia LF, Grosjean J, Joubert M, Darmoni SJ. Medical Informatics. Croatia: InTech; 2012 Mar. Aligning biomedical terminologies in French: towards semantic interoperability in medical applications URL: http://www.chu-rouen.fr/tibs/wp-content/uploads/pdf/Merabti2012.pdf [WebCite Cache ID 6UFCNuqxf]
25. Liu F, Fontelo P. Automatic translations of MEDLINE abstracts using Google Translate. In: AMIA Annu Symp Proc. 2010 Presented at: AMIA Annu Symp; Nov 13-17, 2010; Washington, DC p. 1148.
26. Darmoni SJ, Soualmia LF, Letord C, Jaulent MC, Griffon N, Thirion B, et al. Improving information retrieval using Medical Subject Headings Concepts: a test case on rare and chronic diseases. J Med Libr Assoc 2012 Jul;100(3):176-183 [FREE Full text] [doi: 10.3163/1536-5050.100.3.007] [Medline: 22879806]

Abbreviations

CISMeF: Catalogue et Index des Sites Médicaux de langue Française
LIFO: last in–first out
MeSH: Medical Subject Headings
NLM: National Library of Medicine
PMID: PubMed identifier
TT: transliterated title

Edited by G Eysenbach; submitted 03.09.14; peer-reviewed by H Chneiweiss, Z Lu; comments to author 01.10.14; revised version received 15.10.14; accepted 19.10.14; published 01.12.14.

Please cite as:
Griffon N, Schuers M, Soualmia LF, Grosjean J, Kerdelhué G, Kergourlay I, Dahamna B, Darmoni SJ
A Search Engine to Access PubMed Monolingual Subsets: Proof of Concept and Evaluation in French
J Med Internet Res 2014;16(12):e271
URL: http://www.jmir.org/2014/12/e271/
doi:10.2196/jmir.3836
PMID:25448528

©Nicolas Griffon, Matthieu Schuers, Lina Fatima Soualmia, Julien Grosjean, Gaétan Kerdelhué, Ivan Kergourlay, Badisse Dahamna, Stéfan Jacques Darmoni. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 01.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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

Dr Google and the Consumer: A Qualitative Study Exploring the Navigational Needs and Online Health Information-Seeking Behaviors of Consumers With Chronic Health Conditions

Kenneth Lee[1], BPharm (Hons); Kreshnik Hoti[1], BPharm, PhD; Jeffery David Hughes[1], BPharm, GradDipPharm, MPharm, PhD; Lynne Emmerton[1], BPharm (Hons), PhD

School of Pharmacy, Curtin University, Perth, Australia

Corresponding Author:
Lynne Emmerton, BPharm (Hons), PhD
School of Pharmacy
Curtin University
GPO Box U1987
Perth, 6845
Australia
Phone: 61 8 9266 7352
Fax: 61 8 9266 2769
Email: lynne.emmerton@curtin.edu.au

Abstract

Background: The abundance of health information available online provides consumers with greater access to information pertinent to the management of health conditions. This is particularly important given an increasing drive for consumer-focused health care models globally, especially in the management of chronic health conditions, and in recognition of challenges faced by lay consumers with finding, understanding, and acting on health information sourced online. There is a paucity of literature exploring the navigational needs of consumers with regards to accessing online health information. Further, existing interventions appear to be didactic in nature, and it is unclear whether such interventions appeal to consumers' needs.

Objective: Our goal was to explore the navigational needs of consumers with chronic health conditions in finding online health information within the broader context of consumers' online health information-seeking behaviors. Potential barriers to online navigation were also identified.

Methods: Semistructured interviews were conducted with adult consumers who reported using the Internet for health information and had at least one chronic health condition. Participants were recruited from nine metropolitan community pharmacies within Western Australia, as well as through various media channels. Interviews were audio-recorded, transcribed verbatim, and then imported into QSR NVivo 10. Two established approaches to thematic analysis were adopted. First, a data-driven approach was used to minimize potential bias in analysis and improve construct and criterion validity. A theory-driven approach was subsequently used to confirm themes identified by the former approach and to ensure identified themes were relevant to the objectives. Two levels of analysis were conducted for both data-driven and theory-driven approaches: manifest-level analysis, whereby face-value themes were identified, and latent-level analysis, whereby underlying concepts were identified.

Results: We conducted 17 interviews, with data saturation achieved by the 14th interview. While we identified a broad range of online health information-seeking behaviors, most related to information discussed during consumer-health professional consultations such as looking for information about medication side effects. The barriers we identified included intrinsic barriers, such as limited eHealth literacy, and extrinsic barriers, such as the inconsistency of information between different online sources. The navigational needs of our participants were extrinsic in nature and included health professionals directing consumers to appropriate online resources and better filtering of online health information. Our participants' online health information-seeking behaviors, reported barriers, and navigational needs were underpinned by the themes of trust, patient activation, and relevance.

Conclusions: This study suggests that existing interventions aimed to assist consumers with navigating online health information may not be what consumers want or perceive they need. eHealth literacy and patient activation appear to be prevalent concepts in the context of consumers' online health information-seeking behaviors. Furthermore, the role for health professionals in guiding consumers to quality online health information is highlighted.

XSL•FO
RenderX

(J Med Internet Res 2014;16(12):e262) doi:10.2196/jmir.3706

KEYWORDS

online health information seeking; health information search; health seeking behavior; consumer health information; information needs; Internet; chronic disease; patients; qualitative research; interview

Introduction

The Internet is a popular source of health information among consumers [1-4]. The vastness of health information available online affords consumers unprecedented access to health information. Such access to health information is paramount, given the global push for consumer-focused health care, particularly in the management of chronic health conditions [5-11]. Despite the pervasiveness of the Internet, several of its design features have been identified as potential impediments to consumers' functional access to online health information [12]. As such, navigation of online health information can be problematic. The volume of health information available online can result in information overload [12-14]. The use of overly technical language to convey information, along with the volume of irrelevant content returned from search engine results, the confusing layout of many Web pages, the lack of universal quality requirements for publishing online content, and the abundance of inaccurate or misleading information, are other reported challenges [12]. Despite recognition in the literature of potential navigational issues, to date, little is known about what consumers want in order to better navigate the Internet to access health information.

A comprehensive literature review was conducted to explore interventions aimed at assisting consumers with finding quality online health information related to chronic health conditions [15]. In addition to highlighting the paucity of research in this field, a didactic approach was common among the identified interventions. In these interventions, participants were *taught* how to use the Internet, evaluate the reliability of online health information, or how to use health information databases. Such interventions appear to have been created without research into the navigational needs of consumers.

In order to understand the consumer perspective, we propose the need to explore consumers' online health information-seeking behaviors (HISB). In order to understand consumers' needs for finding online health information, it is essential to know what types of health information are sought online, as well as why, when, where, and how, and the actions that are taken. To date, much of the literature on consumers' online HISB appears to examine only aspects of the concept such as exploring the types of [4,16] and reasons why [16,17] health information is sought, characterizing demographic characteristics of online health information seekers [3,4,18,19], and determining predictors of online health information seeking [1,20]. Furthermore, many studies appear to examine the HISB and online HISB of consumers from populations with particular health conditions [21], such as cancer [17,22-26], spinal cord injury [27,28], and human immunodeficiency virus (HIV) [29-31]. Studies exploring the general online HISB of consumers [3,4,16] included consumers who may not have chronic health conditions. There is thus potential to determine the online HISB and needs of consumers with at least one chronic health condition, representing a significant portion of the adult population.

Studies have suggested the interaction between consumers and health professionals to be a motivator for people to seek health information online [20,21,32-34]. Additionally, consumers' reasons for seeking health information online include enhanced feeling of empowerment [21], perception of improved engagement with their health professionals [21,34], lack of information provided by health professionals [21,35], and dissatisfaction with consumer-health professional interactions [20]. Despite the apparent popularity of the Internet as a source of health information, literature suggests that consumers generally still value the advice given by health professionals above that of health information sourced online [2,3,21,33,36,37]. This suggests that health professionals can play an important role in consumers' online HISB, and we anticipate that consumers' relationships with their health professionals will be pertinent to this investigation.

Further to understanding consumers' navigational needs, a comprehensive understanding of potential barriers to navigation is needed. Recognizing that many consumers have limited ability to find, evaluate, and use health information effectively [38-42], the concept of health literacy is relevant to this investigation. Further to this is consideration of the online environment, with consumers requiring computer and media literacies and skills to overcome the barriers to finding information online, collectively known as eHealth literacy [43]. A more comprehensive picture of the breadth of barriers impeding effective navigation of online health information can thus supplement understanding of navigational needs.

Therefore, this study aims to explore the navigational needs of consumers when searching for health information online for the purpose of self-management of chronic health conditions. To gain a more comprehensive understanding of consumers' needs, this will be explored in the broader context of consumers' online HISB, and potential barriers to navigation will be identified. This approach aims to provide understanding of the navigational needs of health consumers, to inform initiatives and interventions that recognize consumers' needs.

Methods

Procedures and Participants

The apparent need to explore and understand the consumer perspective was approached using qualitative inquiry. Semistructured face-to-face interviews were conducted individually with health consumers. Ethical approval was granted by the Curtin University Human Research Ethics Committee (HR06/2013) and provided for each participant to be offered an AUD $25 gift card to compensate for their travel and time.

XSL•FO
RenderX

Purposive recruitment took place in August-September 2013. Participants were at least 18 years of age, conversant in English, had at least one chronic health condition, and used (or had used) the Internet to find health information related to their condition(s). While an appropriate sample size for qualitative studies should be guided by data saturation, we nominated an initial target of 20 participants, based on commonly cited sample sizes in qualitative studies [44]. The sample was partly sourced by approaching members of the public at nine community pharmacies, with permission of the respective pharmacy store managers and owners, within a practical traveling distance (25 km radius) of the research center in Perth, Western Australia. This coverage was considered to provide a broad demographic of health consumers. This study was also advertised through the student portal on the university website, a local community radio station, and two social media platforms via the university's account.

Interviews

Broad interview questions (Multimedia Appendix 1) surrounding the topics of online HISB and barriers and navigational needs related to finding quality online health information, were devised with reference to the literature and agreed upon by all authors to address the study objectives. Points of interest or points requiring clarification were explored using follow-up questions devised by the interviewer.

Interviews were conducted by the first author and lasted 15-50 minutes. Each interview was audio-recorded with two digital recorders and transcribed verbatim in Microsoft Word as soon as practical after each interview, supplemented by field notes. Grammatical corrections were made to improve flow and readability of the transcripts.

Analysis

Overview

Transcripts were imported into QSR NVivo 10 to facilitate coding and thematic analysis. For analytical rigor, two established approaches to thematic analysis were used. First, we used a data-driven approach (Stage 1 analysis), whereby codes were generated inductively from the data [45,46]. This approach aims to minimize bias in the identification of codes and themes. The codes and themes for this approach were identified by the primary researcher. A subset of transcripts was reviewed by another member of the research team for agreement of codes and themes. On completion of the data-driven analysis, a theory-driven approach was subsequently used by the primary researcher (Stage 2 analysis). This approach required the researchers to formulate a theory about the anticipated results and then develop a provisional list or framework of codes based on the anticipated results [45,46]. Transcripts were then coded using these provisional codes, and additional codes were added if necessary. While this latter approach ensures that the codes and themes identified from the results are focused on the aims of our study, it can restrict the breadth of the findings. Hence, our purpose for using this approach is to confirm the themes identified by the data-driven approach.

Each approach to thematic analysis was further divided into manifest and latent levels of analysis. These two levels reflect how themes are identified. As themes can be broadly defined as patterns found in information, it can range from patterns that are directly observable in the information such as statements made by participants (manifest level), to patterns underlying the information such as potential meaning behind why participants made certain statements (latent level) [45]. We undertook the following specific processes.

Stage 1 Analysis: Data-Driven, Manifest Level

While reading the first interview transcript, codes were identified and created. Codes created from this transcript formed the skeleton of our coding framework, which was continuously revised as more transcripts were coded. Once all transcripts were coded, the codes were grouped into broader categories for identification of themes.

Stage 1 Analysis: Data-Driven, Latent Level

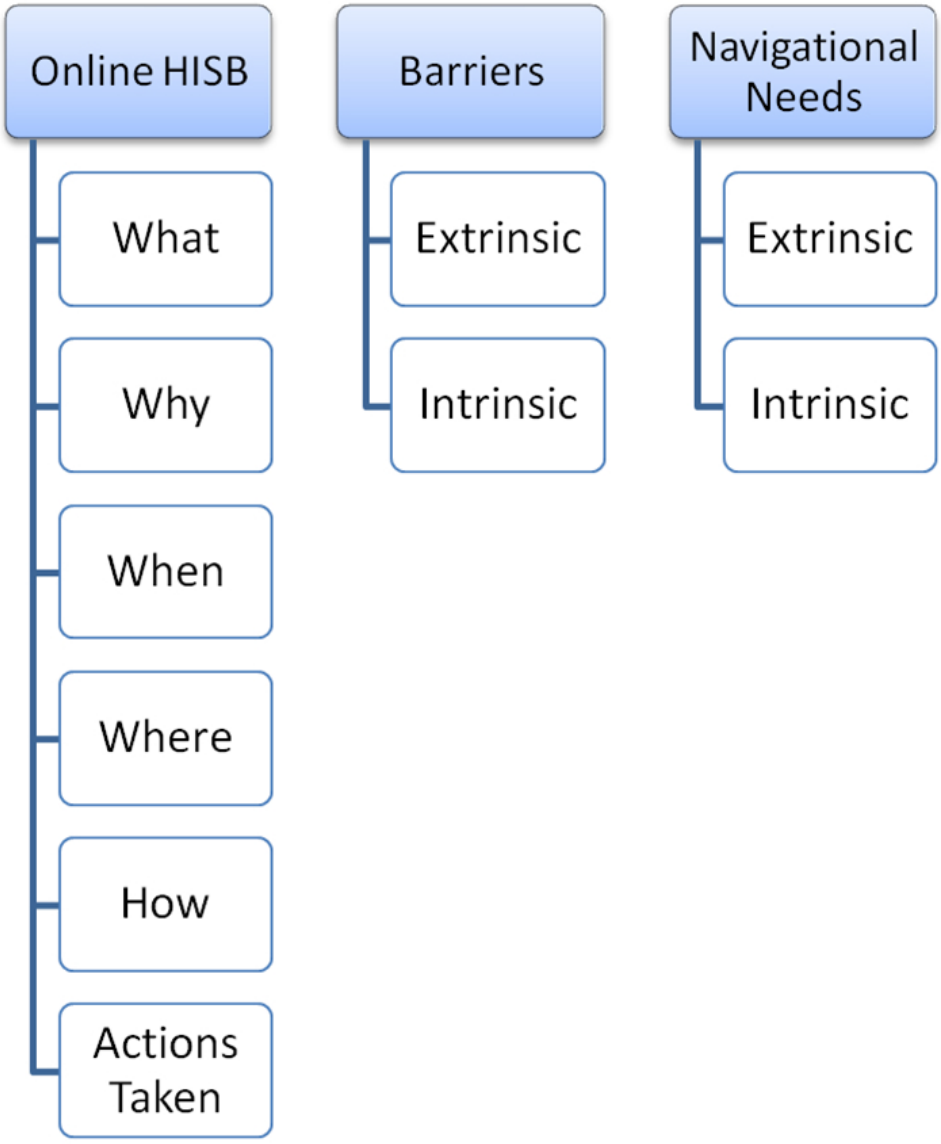
On completion of the manifest-level analysis, the latent-level analysis was conducted. Here, the manifest-level themes identified were continuously compared against all the interview transcripts in attempts to identify underlying patterns (latent themes).

Stage 2 Analyses: Theory-Driven, Manifest and Latent Levels

Figure 1 illustrates the theory that we devised based on our anticipated results. As codes had already been created in Stage 1, a separate QSR NVivo 10 project was established so new codes could be created for this stage. As illustrated in Figure 1, our theory-driven approach categorized online HISB according to *what* types of health information are sought online, *why* health information is sought online, *when* health information is sought online (eg, prior to or after consultation with a health professional), *where* consumers go to obtain online health information (ie, their source(s) of information), *how* consumers go about obtaining online health information (eg, their search strategies), and *actions taken*, that is, what consumers do with the information they find online. We anticipated that barriers to finding online health information could be broadly coded as either extrinsic or intrinsic. For this study, we defined "extrinsic" barriers as those related to the environment or health system. This includes health professionals, health policies, and the design of the Internet/Web pages. We defined "intrinsic" barriers as those pertaining to individual health consumers, such as their motivations to seek health information, and their knowledge medical conditions and their management. These definitions were based on a recent report [47] on health literacy whereby the definition of health literacy was divided into "individual" and "environmental" components. We believe that these two components would apply to our study. We also anticipated that participants' navigational needs would be extrinsic and intrinsic in nature. These categories formed the framework under which subsequent codes were categorized.

XSL•FO
RenderX

Figure 1. Framework for thematic analysis using a theory-driven approach.



Data Saturation

Data saturation was defined by the researchers of this study as the point whereby no new codes could be generated from the data-driven approach. In practice, this point was determined as the participant transcript after which no new codes, relevant to the study, were created in QSR NVivo 10.

Results

Summary

Of the consumers recruited from the nine community pharmacies, 20 initially agreed to participate, of whom 11 attended the interview. Of the remainder, 7 consumers could not be reached, 1 consumer asked to postpone the interview but later retracted due to ailing health, and another did not attend the scheduled interview and could not be contacted thereafter.

After learning of the study via the aforementioned media channels (student portal on the university website, a local community radio station, and two social media platforms via the university's account), 13 consumers enquired about participation. Of these, 8 people agreed to participate, but 1 did not attend the scheduled interview and could not be reached thereafter.

In total, 17 eligible participants completed the interview: 10 participants recruited from pharmacies, 4 via community radio, and 3 via university advertising channels. An additional recruit was excluded due to ineligibility identified during the interview. There were 9 female and 8 male participants. The age of participants ranged from 19-85 years, with the most common age category being 50-60 years. The majority of participants appeared to be native English speakers. Participants comprised 3 university students, 8 in the workforce, and 6 retirees. The participants' diagnosed medical conditions were not formally recorded for privacy reasons; however, their conditions revealed during the interviews were broad ranging, and a majority of participants reported having more than one chronic health condition.

Data saturation was perceived by the 14th interview; no new codes relevant to this study were generated from the remaining three interview transcripts.

XSL•FO
RenderX

Manifest-Level Analysis

Tables 1 and 2 reveal a wide range of responses related to *what* types of online health information were sought, *why*, *when*, *where*, and *how* online information was sought, and what consumers reported doing with the health information they found online (*actions taken*).

As indicated in Table 1, participants most commonly reported seeking health information related to understanding their medical conditions and the medications prescribed by their health professionals (*what*). Common reasons for seeking information online were to be more informed about their health and were commonly related to consultations with their health professionals (*why*). The most common time in which participants were more likely to seek information (*when*) was after consulting a health professional.

As indicated in Table 2, despite typically searching for health information after consulting a health professional, the most common approach was the use of a search engine (*where* and *how*). Once participants obtained health information, their actions appeared to largely revolve around the consultations with their health professionals in the sense that participants either decided to relay their findings to their health professionals or make decisions about whether to consult their health professionals for advice or medical attention (*actions taken*). Some participants also reported that the information sourced online assisted in making decisions about whether to use therapeutic products or to trial lifestyle modifications to supplement advice given by their health professionals.

Table 1. Online HISB: what, why, when (responses are listed in approximate order of frequency from most to least frequently mentioned).

What	Why	When
Medicines or medical devices—includes side effects, and indications for the products	To be more informed	Following a consultation with a health professional
Medical conditions	To clarify/verify information discussed during consultation	Before a consultation with a health professional
Lifestyle information (eg, information on diets and exercise)	Because the Internet is accessible	When required
Information about individual health professionals, medical clinics and hospitals (eg, appropriate specialists for particular conditions)	Emotional support, eg, to read about experiences of others with the same condition(s)	Before and after a consultation
Natural products	Out of interest	
Information about disease-specific associations, eg, Cancer Council and Diabetes Australia	Disagreement with points made by a health professional	
Medical glossary	To seek alternative/additional treatment options	
	Insufficient information provided during a consultation	
	Urgency to know	
	To take charge of one's life	
	Self-management of a perceived minor condition, eg, common cold	
	Limited time during consultation	
	To have written information to read	
	Infrequency of interaction with health professional	

XSL•FO
RenderX

Table 2. Online HISB: where, how, actions taken (responses are listed in approximate order of frequency from most to least frequently mentioned).

Where	How	Actions taken
Search engines (eg, Google, Bing)	Start with results from a search engine	Discuss information with a health professional
Disease-specific association websites	Direct URL	Decide whether to consult with a health professional
Forums/support groups	Start with websites recommended by others	Decide whether to purchase/use a medication or natural product
Wikipedia		Trial lifestyle modifications
General health websites (eg, Better Health Channel)		
Website recommended by health professionals		
Research databases		
eNewsletters		
Private health insurer websites (eg, Medibank)		

Table 3 summarizes participants' self-reported barriers to online HISB. Table 4 summarizes participants' reported navigational needs for finding online health information.

Participants' self-reported barriers to online HISB (Table 3) were able to be categorized into "extrinsic" and "intrinsic" subcategories. This suggests that our findings are aligned with our theorized framework illustrated in Figure 1. Extrinsic barriers included issues with the information or presentation of information online, the volume of information available online, and issues with the relationship between consumers and their health professionals. Regarding intrinsic barriers, limited "eHealth literacy skills" was the prevalent theme.

When participants were asked about their navigational needs (Table 4), their comments focused on extrinsic needs and encompassed improving access to, or transparency of, quality information, reducing/blocking access to poor quality information, and involving health professionals in guiding consumers to quality health information. No intrinsic navigational needs were reported by our participants.

Table 3. Participants' self-reported barriers to online HISB.

Category	Feedback
Extrinsic	Availability/accessibility of content, eg, difficulty in accessing content published in research journals
	Use of medical jargon
	Inconsistency of information across different sources
	Volume of information available
	Poor interaction/relationship with health professional, eg, low level of trust in the advice given by health professionals
Intrinsic	Limited eHealth literacy skills
	Limited knowledge of credible websites
	Unsure of information need
	Limited time available to search for information
	Lack of motivation to search for information

Table 4. Participants' self-reported navigational needs.

Category	Feedback
Extrinsic	Greater availability/accessibility of content
	Single/designated destination for credible information
	Health professionals' input, such as providing an annotated list of potentially useful websites
	Blocking untrustworthy websites
	Stricter rules for publishing information online
	Improved webpage layout and features
Intrinsic	None reported by participants

XSL•FO
RenderX

Latent-Level Analysis

Our latent-level analysis revealed that there were three themes underlying one or more aspects of our participants' online HISB, and barriers and navigational needs for finding online health information: trust, patient activation, and relevance. In particular, these three themes appear to provide potential reasons as to why participants chose certain sources of information and why they were motivated to search for health information online, and they reinforce the important role online health information plays in the self-management of chronic health conditions.

Trust

Throughout the interviews, the idea of trust became evident as a factor associated with online HISB. This idea of trust encompassed both trust in health professionals, and trust in certain online sources such as search engines. Trust in advice given by health professionals is exemplified in the following quotations:

> *If it's something that the doctor has prescribed for me, generally I don't research that, because I think the doctor has decided that "that's what I want", and he decided what the dose is and told me what the side effects might be—so I don't actually go into that.* [P7]

> *The Internet is part of my life medically, but only a relatively small part. I rely mostly on my GP and specialist for most things.* [P12]

> *Yes, and [my doctor is] very good with latest research and things. He will say, "Oh I read this, and that doesn't work anymore". What was that thing that I was going to talk to him about next time that's been recently in the TV—is it Statins for the high blood pressure? That's one thing I will ask him. So I'd rather trust him than say, 60 Minutes [television show].* [P17]

Despite this trust in their health professionals, participants still searched for health information online. This suggests that the idea of trust also applied to online sources. From one participant reflecting on Wikipedia:

> *Generally, I'd say I trust it, but I know it can be very variable because it's written by all sorts of people. They do have a team of writers, but they are, as the name implies, almost anyone can add to it. So you've got to look carefully at the qualifications of who's written that particular piece.* [P7]

Further to the topic of trust in online sources, search engines such as Google were mentioned by many participants as a trusted source. The following two examples suggest that the participants trust the Google search engine to provide them with a starting point to finding relevant information.

When asked how online information is sought:

> *For example, for glandular fever, I will put in "recurring glandular fever". When you do it in Google, it virtually thinks ahead for you. So it's fantastic in that way because maybe you haven't thought of exactly the right words and it comes up with half a dozen—it triggers off this set of words and then you just click on the one that's closest to the one that you're thinking.* [P11]

When asked how online information is sought:

> *In general terms, I start off with Google and see where that takes me.* [P13]

However, while participants generally appeared to trust certain online sources of health information, the majority of participants still sought clarification with health professionals for their discipline expertise. The following participant mentioned seeking clarification with their doctor after obtaining health information online:

> *I had an issue last year where my feet were swelling and I was getting a bit of a pigmented pattern. I looked that up to see what it was…I thought "Oh, should I be worried? What's this? I should go mention it to my doctor"…It prompted me to say "I better go and ask a question about it next time I see him [the doctor]".* [P4]

A lack of trust appears to contribute to some of the identified barriers in Table 3. For example, individuals may not trust online health information because they are unable to appraise the information for its reliability. This could be due to an individual's limited eHealth literacy, among other factors.

When asked for his/her opinion about health information found online:

> *I don't put a great deal of trust into it because, you know, so many different things come on the Internet. How do you weigh the worth against another when it's not your field?* [P8]

Regarding navigational needs, the theme of trust appeared to be a factor influencing participants' comments. For example, one participant commented that their doctor provided an annotated list of relevant websites, which the participant found valuable:

> *He actually gave me a list of websites that are good ones to use...I like the fact that I was given a choice, given a list, so I could go through them all and choose one for me. I wasn't just given one and told "this will suit all your needs"—because it's never really going to and I guess in that sense, I did go through and decide to pick the one that I liked. But you know, you could pick and choose and compare.* [P15]

While many participants reported that they trust a variety of online information sources, most of these participants also reported that they trust one or more of their health professionals. This theme of trust appeared to underlie participants' online HISB and barriers and navigational needs for finding online health information. Interestingly, it appears that trust in health professionals prevailed over trust in online sources.

Patient Activation

Patients or health consumers are considered to be activated if they "believe patients have important roles to play in self-managing care, collaborating with providers, and maintaining their health. They know how to manage their

XSL•FO
RenderX

condition and maintain functioning and prevent health declines; and they have the skills and behavioral repertoire to manage their condition, collaborate with their health providers, maintain their health functioning, and access appropriate and high-quality care" [48].

The theme of "patient activation" aligned primarily with the reasons for consumers to seek online health information: an urgency to know, to be more informed, to find alternative treatment options, to take charge of one's life, and to manage self-perceived minor conditions. Components of this definition were reflected in comments from participants.

When asked why the participant sought information online:

> *Just to look for information that my doctor wasn't able to tell me…They didn't know what the cause was. So I was really adamant to find out what the possible cause would be.* [P1]

When asked about whether the participant experiences difficulty finding health information online:

> *I generally find that I can find what I want pretty quickly. As I say, there's all sorts of things with Google. It's quite massive the amount of information that is there, but you've got to use a little bit of intelligence to get through it.* [P2]

On the topic of why health information is sought online as opposed to seeking a health professional:

> *There's sometimes a waitlist…if I have to wait a week to see a doctor, I lose interest, so I kind of go, "ok, I'm going to research this", do it and that's it.* [P3]

When asked why the participant sought information online:

> *More so clarification and to get information…because I want to be informed… I don't want to be a passive patient…I want to know what's going on, I want to make informed decisions. So if they decide on a particular course of treatment, I want to know what's going on and if there is an alternative.* [P16]

In terms of patient activation influencing barriers to finding online health information, one participant claimed:

> *I'm not terribly patient of the Internet. I would turn it off and ask someone. I think some people will sit there because they're happy looking at a screen and clicking forever, whereas if I can't get something reasonably quickly, or know where to start, I'll just get a book off the shelf or see someone.* [P6]

This appears to reflect the participant's lack of motivation to use the Internet to find health information. However, it is important to note that this participant still appeared to demonstrate a degree of activation from being motivated to seek health information from other sources.

While participants highlighted their desire for health professionals to play a greater role in assisting their navigation of online health information (Table 4), many participants also expressed the need to improve various aspects of the Internet, such as blocking untrustworthy websites. This reinforces the desire for participants to use the Internet to find health information and suggests patient activation is a potential factor having an impact on their navigational needs.

Relevance

While some participants reported that they also looked online for information about health topics that were of interest but were unrelated to their health issues, the primary reason for searching online was to find health information relevant to their condition(s). This suggests that our participants appeared to value the role online health information can play in self-management of chronic health conditions.

When asked what types of health information is sought:

> *Well, usually it was for pain patches, because I want to investigate about putting on weight with them, because I have a lot of fluid retention…So I want to look at that and see.* [P5]

When asked what the participant would do regarding an information need:

> *I would look at the Internet a number of times, because there's so much conflicting information out there. And then I would try and keep narrowing it down to things that may be helpful for my condition.* [P15]

In terms of barriers to finding online health information, the volume of irrelevant information can impede access to relevant online health information. One participant commented that this can be particularly important when one has comorbidities:

> *There's so much information out there. It's having more than one medical condition, for me, makes it complicating. For example, I'll be looking up depression and diabetes, and there's not always a cross-match to compare them.* [P3]

The navigational needs listed in Table 4 suggest participants' desire to be more able to find relevant and reliable online health information, as illustrated by the following participant:

> *I'd love to have a search function that can say "I want to hear about blah, blah and blah, but not blah".* [P17]

Discussion

Principal Findings

Our study suggests that the didactic emphasis of existing interventions is not in line with consumers' needs when searching for online health information. Our participants' navigational needs were extrinsic in nature. They wanted to be able to better navigate online health information by improving systems supporting their health information-seeking activities. These systems include design features of the Internet that have been reported [12] as potential impediments to functional access, as well as health professionals guiding consumers to appropriate online resources. This finding reinforces the need for greater efforts into addressing such design issues of the Internet, as well as suggesting that health professionals can play a role in consumers' navigation of online health information.

XSL•FO
RenderX

Regarding the online HISB of consumers with chronic health conditions, we report that such behaviors commonly revolve around (either before, or more commonly after) consultations with health professionals. Various reported reasons *why* our participants sought information (Table 1) and *actions taken* (Table 2) by our participants appear to be in line with a study of consumers' online HISB in the context of medical consultations [21,34]. Given the long-term management requirements for chronic health conditions and the need for regular appointments with health professionals, it appears reasonable that online HISB is aligned with health professional consultations. This finding reinforces the importance of health professionals in influencing the online HISB of consumers with chronic health conditions and further supports literature reporting consumers' trust in health professionals with regards to health information [2,3,33,36,37].

Regarding *how* online health information is found (Table 2), search engines were often the first port of call. This finding confirms other studies whereby search engines have been identified as a dominant source of health information [16,49-51]. Our findings suggest that this could be attributed to identified intrinsic barriers such as a lack of awareness of relevant and reliable websites. However, further studies are recommended to determine other potential reasons. Given the prevalent use of search engines to acquire health information and some participants' expressed "trust in Google" to provide relevant results, it appears that search engines can also play an important role in influencing consumers' online HISB. This is reinforced by the themes of trust and relevance, which we identified as two of the three themes underlying consumers' online HISB and navigational needs.

In terms of the extrinsic barriers (Table 3), our findings suggest that greater involvement by health professionals could assist in minimizing such barriers. For instance, by providing consumers with greater guidance on where to go for quality resources that are pertinent and comprehensible to the individual consumer, the issues of medical jargon, volume of information, and inconsistency of information across different sources may be addressed. In doing so, this could also perceivably improve the health professional–consumer relationship. Such involvement is consistent with one or more of participants' navigational needs (Table 4). Similarly, given the prominent use of search engines in the context of online HISB and consumers' navigational needs, improving filtering of poor quality websites by search engines could help address one or more of the identified barriers (Table 3) and is also in line with participants' navigational needs (Table 4). These suggestions require further investigation to ensure validity in our claims.

Our study identified eHealth literacy skills as a prominent intrinsic barrier to finding online health information. This prevalence is suggestive of the need to address issues of Health literacy and eHealth literacy, as suggested by the multitude of studies conducted in these fields [38-41,43]. We also identified that the theme of patient activation underlies consumers' online HISB. Given the prominence of eHealth literacy as a barrier, as well as the importance of patient activation in consumers' online HISB, there may be associations between consumers' online HISB, eHealth literacy and patient activation, and their role in self-management of chronic health conditions. Further investigation is required to explore the possibility of such associations.

The themes of trust, patient activation, and relevance, identified via latent-level analysis, underpinned one or more aspects of our participants' online HISB and barriers and navigational needs for finding online health information. While the theme of trust has been identified in various studies as a predictor of online HISB [1,20,32,52], the concept of patient activation appears less studied in the online health information context. We propose that patient activation will remain crucial in light of social trends towards self-care. One would expect that consumers with chronic conditions would most commonly search for information relating to their own conditions. However, it is noteworthy when considering the plethora of online information and potential challenges for consumers in identifying relevance of information. While further studies are required to examine the significance and interrelationships of these three themes, we have provided a framework for future initiatives to support consumers in their navigation of online health information.

Strengths and Limitations

A key strength of this study lies in our use of two approaches to analysis (data-driven and theory-driven analyses) to assist with consolidation of the themes. Further to minimizing bias in analysis, as mentioned above, the use of a data-driven approach has been suggested to offer greater criterion and construct validity over a theory-driven approach [45]. A limitation of the data-driven approach is potential identification of codes and themes that may not be relevant to the study objectives [45]. We believe that combining these two approaches improves the rigor of our analysis by mitigating limitations from each individual approach. As such, we believe our findings can be considered more robust compared to using each analytical approach separately.

By separating manifest and latent-level themes, and by demonstrating what we classified as manifest and latent-level themes, we have added transparency to our methods of analysis. This facilitates peer critique of our interpretation and assessment of the robustness of our findings.

While many studies focused on HISB and online HISB in the context of specific chronic conditions [17,22-31,53-55], our study qualitatively explored online HISB for a variety of chronic health conditions. Despite the broader focus of our study, manifest level themes identified (Table 1) including the types of and reasons why online health information is sought, is in line with findings from various disease-specific studies. This corroboration of findings, despite differing research designs, settings, and sampling approaches, suggests consistency of our interpretation of our results and robustness in our findings. Our study therefore provides a foundation to better understanding the overall concept of online HISB.

At the time of writing, we had not identified any other study that attempted to identify consumers' navigational needs when searching for online health information. Hence this study provides the foundations for investigation into the navigational

XSL•FO
RenderX

needs of health consumers in the context of online health information seeking.

While various means of recruiting participants were used, all participants were from the metropolitan area in Perth, Western Australia. This limits the generalizability of our findings. Further, the nature of qualitative research means that generalization of findings is not appropriate. Nevertheless, the diversity of our participants' backgrounds, the number and types of chronic health conditions of our participants, and the match between our prominent manifest level themes with other studies suggest that our findings could be transferable to a wider population of consumers with chronic health conditions who use the Internet to find health information. Quantitative studies may be a useful supplement to our qualitative data.

A further limitation to our study is that our qualitative approach does not allow statistical analysis of interrelationships between manifest-level themes. Our data suggest potential relationships between various manifest-level themes and latent-level themes. A quantitative study is recommended for confirmation.

Further Research

Our research uniquely explores consumers' perspectives on how they would like to better navigate the Internet to find online health information, for the purpose of self-management of chronic health conditions. A comprehensive literature review highlighted that previous studies focused on delivering educational-type interventions [15]; however, our study reveals that consumers may prefer other avenues. Given that this, to our knowledge, is the first study inviting the consumer perspective, we recommend further investigation into consumers' navigational needs for online health information, to inform practical solutions.

Further, the prominence of the concepts of online HISB, eHealth literacy, and patient activation in our study warrants further exploration. Other studies have examined the issue of eHealth literacy alone [43,56-58], and one study examined the interplay between health literacy and patient activation [59]. A more comprehensive understanding of these three concepts could further assist in addressing the navigational needs of consumers.

Due to the qualitative data, statistically significant correlations between our identified themes cannot be examined. Quantitative research is in progress by the researchers of this study to examine potential correlations between the prominent themes identified in this study.

The navigational needs for finding online health information, as recommended by participants, appear to address only the extrinsic barriers identified in this study. Future research could investigate reasons for this finding.

Conclusions

This study explores consumers' navigational needs for finding online health information and identifies various barriers to consumers' online HISB. In doing so, it highlights that consumers may desire interventions other than the educational-type interventions provided to date and highlights avenues for future research. eHealth literacy and patient activation were prominent themes in the context of online HISB and are possibly related.

Of particular note, our findings suggest consumers' desire for health professionals to play a role in guiding them to find relevant and reliable online health information. Consumers will continue to source health information online, and those motivated to source such information require assurance of its accuracy. By providing an understanding of how consumers can potentially be assisted in finding quality online health information, our study serves as a foundation towards assisting consumers with better self-management of their chronic health conditions.

Acknowledgments

The researchers of this study would like to acknowledge all the consumers who participated in the interviews and the pharmacy owners and managers for permitting recruitment to take place at their premises. KL is supported by an Australian Postgraduate Award.

Authors' Contributions

KL conceived and designed the study with assistance from LE, KH and JH. KL conducted the recruitment, interviewing, transcription, coding, analysis, and drafted the manuscript. LE reviewed a sample of the transcripts to verify the interpretation of identified codes and themes. All authors were involved in reviewing and revising the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Interview guide.

[PDF File (Adobe PDF File), 11KB - jmir_v16i12e262_app1.pdf]

References

XSL•FO
RenderX

1. Miller LM, Bell RA. Online health information seeking: the influence of age, information trustworthiness, and search challenges. J Aging Health 2012 Apr;24(3):525-541. [doi: 10.1177/0898264311428167] [Medline: 22187092]
2. Andreassen HK, Bujnowska-Fedak MM, Chronaki CE, Dumitru RC, Pudule I, Santana S, et al. European citizens' use of E-health services: a study of seven countries. BMC Public Health 2007;7:53 [FREE Full text] [doi: 10.1186/1471-2458-7-53] [Medline: 17425798]
3. Fox S, Jones S. The Social Life of Health Information.: Pew Internet & American Life Project; 2009. URL: http://www.pewinternet.org/2009/06/11/the-social-life-of-health-information/ [accessed 2014-07-12] [WebCite Cache ID 6R0Q6Jqpa]
4. Fox S. Health Topics.: Pew Internet & American Life Project; 2011. URL: http://www.pewinternet.org/2011/02/01/health-topics-2/ [accessed 2014-09-23] [WebCite Cache ID 6SnedkDik]
5. Aujoulat I, d'Hoore W, Deccache A. Patient empowerment in theory and practice: polysemy or cacophony? Patient Educ Couns 2007 Apr;66(1):13-20. [doi: 10.1016/j.pec.2006.09.008] [Medline: 17084059]
6. Barlow J, Wright C, Sheasby J, Turner A, Hainsworth J. Self-management approaches for people with chronic conditions: a review. Patient Educ Couns 2002;48(2):177-187. [Medline: 12401421]
7. Bergeson SC, Dean JD. A systems approach to patient-centered care. JAMA 2006 Dec 20;296(23):2848-2851. [doi: 10.1001/jama.296.23.2848] [Medline: 17179462]
8. Bertakis KD, Azari R. Determinants and outcomes of patient-centered care. Patient Educ Couns 2011 Oct;85(1):46-52. [doi: 10.1016/j.pec.2010.08.001] [Medline: 20801601]
9. Breen GM, Wan TT, Zhang NJ, Marathe SS, Seblega BK, Paek SC. Improving doctor-patient communication: examining innovative modalities vis-a-vis effective patient-centric care management technology. J Med Syst 2009 Apr;33(2):155-162. [Medline: 19397100]
10. Kjeldmand D, Holmström I, Rosenqvist U. How patient-centred am I? A new method to measure physicians' patient-centredness. Patient Educ Couns 2006 Jul;62(1):31-37. [doi: 10.1016/j.pec.2005.05.017] [Medline: 16024208]
11. McNutt RA. Shared medical decision making: problems, process, progress. JAMA 2004 Nov 24;292(20):2516-2518. [doi: 10.1001/jama.292.20.2516] [Medline: 15562133]
12. Cline RJ, Haynes KM. Consumer health information seeking on the Internet: the state of the art. Health Educ Res 2001 Dec;16(6):671-692 [FREE Full text] [Medline: 11780707]
13. McGrath I. Information superhighway or information traffic jam for healthcare consumers? Clin Perform Qual Health Care 1997;5(2):90-93. [Medline: 10167219]
14. Metzger MJ, Flanagin AJ. Using Web 2.0 technologies to enhance evidence-based medical information. J Health Commun 2011;16 Suppl 1:45-58. [doi: 10.1080/10810730.2011.589881] [Medline: 21843095]
15. Lee K, Hoti K, Hughes JD, Emmerton LM. Interventions to assist health consumers to find reliable online health information: a comprehensive review. PLoS One 2014;9(4):e94186 [FREE Full text] [doi: 10.1371/journal.pone.0094186] [Medline: 24710348]
16. McDaid D, Park AL. Online Health: Untangling the Web. 2010. URL: https://www.bupa.com.au/staticfiles/Bupa/HealthAndWellness/MediaFiles/PDF/LSE_Report_Online_Health.pdf [accessed 2014-09-23] [WebCite Cache ID 6Snf9eDha]
17. Lee SY, Hawkins R. Why do patients seek an alternative channel? The effects of unmet needs on patients' health-related Internet use. J Health Commun 2010 Mar;15(2):152-166. [doi: 10.1080/10810730903528033] [Medline: 20390984]
18. Yan YY. Online health information seeking behavior in Hong Kong: an exploratory study. J Med Syst 2010 Apr;34(2):147-153. [Medline: 20433053]
19. Cotten SR, Gupta SS. Characteristics of online and offline health information seekers and factors that discriminate between them. Soc Sci Med 2004 Nov;59(9):1795-1806. [doi: 10.1016/j.socscimed.2004.02.020] [Medline: 15312915]
20. Tustin N. The role of patient satisfaction in online health information seeking. J Health Commun 2010 Jan;15(1):3-17. [doi: 10.1080/10810730903465491] [Medline: 20390974]
21. Caiata-Zufferey M, Abraham A, Sommerhalder K, Schulz PJ. Online health information seeking in the context of the medical consultation in Switzerland. Qual Health Res 2010 Aug;20(8):1050-1061. [doi: 10.1177/1049732310368404] [Medline: 20442347]
22. Roberts KJ, Lepore SJ, Urken ML. Quality of life after thyroid cancer: an assessment of patient needs and preferences for information and support. J Cancer Educ 2008;23(3):186-191. [doi: 10.1080/08858190802247762] [Medline: 18709591]
23. McCaughan E, McKenna H. Information-seeking behaviour of men newly diagnosed with cancer: a qualitative study. J Clin Nurs 2007 Nov;16(11):2105-2113. [doi: 10.1111/j.1365-2702.2006.01785.x] [Medline: 17331096]
24. Chou WY, Liu B, Post S, Hesse B. Health-related Internet use among cancer survivors: data from the Health Information National Trends Survey, 2003-2008. J Cancer Surviv 2011 Sep;5(3):263-270. [doi: 10.1007/s11764-011-0179-5] [Medline: 21505861]
25. James N, Daniels H, Rahman R, McConkey C, Derry J, Young A. A study of information seeking by cancer patients and their carers. Clin Oncol (R Coll Radiol) 2007 Jun;19(5):356-362. [doi: 10.1016/j.clon.2007.02.005] [Medline: 17399963]
26. Noh HI, Lee JM, Yun YH, Park SY, Bae DS, Nam JH, et al. Cervical cancer patient information-seeking behaviors, information needs, and information sources in South Korea. Support Care Cancer 2009 Oct;17(10):1277-1283. [doi: 10.1007/s00520-009-0581-y] [Medline: 19172301]

XSL•FO
RenderX

27. Burkell JA, Wolfe DL, Potter PJ, Jutai JW. Information needs and information sources of individuals living with spinal cord injury. Health Info Libr J 2006 Dec;23(4):257-265. [doi: 10.1111/j.1471-1842.2006.00686.x] [Medline: 17177946]
28. Matter B, Feinberg M, Schomer K, Harniss M, Brown P, Johnson K. Information needs of people with spinal cord injuries. J Spinal Cord Med 2009;32(5):545-554 [FREE Full text] [Medline: 20025150]
29. Samal L, Saha S, Chander G, Korthuis PT, Sharma RK, Sharp V, et al. Internet health information seeking behavior and antiretroviral adherence in persons living with HIV/AIDS. AIDS Patient Care STDS 2011 Jul;25(7):445-449 [FREE Full text] [doi: 10.1089/apc.2011.0027] [Medline: 21682586]
30. O'Grady L. Meeting health information needs of people with HIV/AIDS: sources and means of collaboration. Health Info Libr J 2008 Dec;25(4):261-269. [doi: 10.1111/j.1471-1842.2007.00764.x] [Medline: 19076672]
31. Kalichman SC, Benotsch EG, Weinhardt L, Austin J, Luke W, Cherry C. Health-related Internet use, coping, social support, and health indicators in people living with HIV/AIDS: preliminary results from a community survey. Health Psychol 2003 Jan;22(1):111-116. [Medline: 12558209]
32. Hou J, Shim M. The role of provider-patient communication and trust in online sources in Internet use for health-related activities. J Health Commun 2010;15 Suppl 3:186-199. [doi: 10.1080/10810730.2010.522691] [Medline: 21154093]
33. Murray E, Lo B, Pollack L, Donelan K, Catania J, White M, et al. The impact of health information on the internet on the physician-patient relationship: patient perceptions. Arch Intern Med 2003 Jul 28;163(14):1727-1734. [doi: 10.1001/archinte.163.14.1727] [Medline: 12885689]
34. Attfield SJ, Adams A, Blandford A. Patient information needs: pre- and post-consultation. Health Informatics J 2006 Jun;12(2):165-177. [doi: 10.1177/1460458206063811] [Medline: 17023406]
35. Rozmovits L, Ziebland S. What do patients with prostate or breast cancer want from an Internet site? A qualitative study of information needs. Patient Educ Couns 2004 Apr;53(1):57-64. [doi: 10.1016/S0738-3991(03)00116-2] [Medline: 15062905]
36. Fox S. After Dr Google: peer-to-peer health care. Pediatrics 2013 Jun;131 Suppl 4:S224-S225 [FREE Full text] [doi: 10.1542/peds.2012-3786K] [Medline: 23729765]
37. Hesse BW, Moser RP, Rutten LJ. Surveys of physicians and electronic health information. N Engl J Med 2010 Mar 4;362(9):859-860. [doi: 10.1056/NEJMc0909595] [Medline: 20200398]
38. 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]
39. Cutilli CC, Bennett IM. Understanding the health literacy of America: results of the National Assessment of Adult Literacy. Orthop Nurs 2009;28(1):27-32; quiz 33 [FREE Full text] [doi: 10.1097/01.NOR.0000345852.22122.d6] [Medline: 19190475]
40. Manning DL, Dickens C. Health literacy: more choice, but do cancer patients have the skills to decide? Eur J Cancer Care (Engl) 2006 Dec;15(5):448-452. [doi: 10.1111/j.1365-2354.2006.00687.x] [Medline: 17177901]
41. Nielsen-Bohlman L, Panzer AM, Kindig DA. Health literacy: a prescription to end confusion. Washington, DC: National Academies Press; 2004.
42. World Health Organization. Health literacy and health behavior. 2011. URL: http://www.who.int/healthpromotion/conferences/7gchp/track2/en/ [accessed 2014-07-12] [WebCite Cache ID 6R0P5U0gK]
43. Norman CD, Skinner HA. eHealth Literacy: Essential Skills for Consumer Health in a Networked World. J Med Internet Res 2006;8(2):e9 [FREE Full text] [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]
44. Mason M. Sample size and saturation in PhD studies using qualitative interviews. Forum Qualitative Social Research 2010 Sep;11(3):- [FREE Full text]
45. Boyatzis RE. Transforming qualitative information: thematic analysis and code development. Thousand Oaks, CA: Sage Publications; 1998.
46. Miles MB, Huberman AM, Saldaña J. Qualitative data analysis: a methods sourcebook. 3rd edition. Thousand Oaks, CA: Sage Publications; 2013.
47. Australian Commission on Safety and Quality in Health Care. Health literacy: Taking action to improve safety and quality. Sydney (Australia): ACSQHC; 2014. URL: http://www.safetyandquality.gov.au/wp-content/uploads/2014/08/Health-Literacy-Taking-action-to-improve-safety-and-quality.pdf [accessed 2014-10-06] [WebCite Cache ID 6T7I4e39T]
48. Hibbard JH, Stockard J, Mahoney ER, Tusler M. Development of the Patient Activation Measure (PAM): conceptualizing and measuring activation in patients and consumers. Health Serv Res 2004 Aug;39(4 Pt 1):1005-1026 [FREE Full text] [doi: 10.1111/j.1475-6773.2004.00269.x] [Medline: 15230939]
49. Eysenbach G, Köhler C. How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. BMJ 2002 Mar 9;324(7337):573-577 [FREE Full text] [Medline: 11884321]
50. Fox S. Online Health Search.: Pew Internet & American Life Project; 2006. URL: http://www.pewinternet.org/files/old-media/Files/Reports/2006/PIP_Online_Health_2006.pdf.pdf [accessed 2014-07-12] [WebCite Cache ID 6R0OwsHEC]
51. Geana MV, Daley CM, Nazir N, Cully L, Etheridge J, Bledowski C, et al. Use of online health information resources by American Indians and Alaska Natives. J Health Commun 2012 Aug;17(7):820-835 [FREE Full text] [doi: 10.1080/10810730.2011.650831] [Medline: 22642739]

XSL•FO
RenderX

52. AlGhamdi KM, Moussa NA. Internet use by the public to search for health-related information. Int J Med Inform 2012 Jun;81(6):363-373. [doi: 10.1016/j.ijmedinf.2011.12.004] [Medline: 22217800]
53. Raynor DK, Savage I, Knapp P, Henley J. We are the experts: people with asthma talk about their medicine information needs. Patient Educ Couns 2004 May;53(2):167-174. [doi: 10.1016/S0738-3991(03)00126-5] [Medline: 15140456]
54. Rezailashkajani M, Roshandel D, Ansari S, Zali MR. Knowledge of disease and health information needs of the patients with inflammatory bowel disease in a developing country. Int J Colorectal Dis 2006 Jul;21(5):433-440. [doi: 10.1007/s00384-005-0030-4] [Medline: 16132999]
55. Rodgers S, Dyas J, Molyneux AW, Ward MJ, Revill SM. Evaluation of the information needs of patients with chronic obstructive pulmonary disease following pulmonary rehabilitation: a focus group study. Chron Respir Dis 2007;4(4):195-203. [doi: 10.1177/1479972307080698] [Medline: 18029432]
56. 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;15(5):e93 [FREE Full text] [doi: 10.2196/jmir.2645] [Medline: 23639979]
57. Manafò E, Wong S. Assessing the eHealth literacy skills of older adults: A preliminary study. J Consum Health Internet 2012 Oct;16(4):369-381. [doi: 10.1080/15398285.2012.701163]
58. Norman C. eHealth literacy 2.0: problems and opportunities with an evolving concept. J Med Internet Res 2011;13(4):e125 [FREE Full text] [doi: 10.2196/jmir.2035] [Medline: 22193243]
59. Smith SG, Curtis LM, Wardle J, von Wagner C, Wolf MS. Skill set or mind set? Associations between health literacy, patient activation and health. PLoS One 2013;8(9):e74373 [FREE Full text] [doi: 10.1371/journal.pone.0074373] [Medline: 24023942]

Abbreviations

HISB: health information-seeking behaviors
HIV: human immunodeficiency virus

Edited by G Eysenbach; submitted 15.07.14; peer-reviewed by S Koch-Weser, B St. Jean; comments to author 14.08.14; revised version received 24.09.14; accepted 20.10.14; published 02.12.14.

Please cite as:
Lee K, Hoti K, Hughes JD, Emmerton L
Dr Google and the Consumer: A Qualitative Study Exploring the Navigational Needs and Online Health Information-Seeking Behaviors of Consumers With Chronic Health Conditions
J Med Internet Res 2014;16(12):e262
URL: http://www.jmir.org/2014/12/e262/
doi:10.2196/jmir.3706
PMID:25470306

©Kenneth Lee, Kreshnik Hoti, Jeffery David Hughes, Lynne Emmerton. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 02.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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 Effect of Individual Factors on Health Behaviors Among College Students: The Mediating Effects of eHealth Literacy

WanChen Hsu[1], PhD; ChiaHsun Chiang[2]; ShuChing Yang[2], PhD

[1]Center for Teaching & Learning Development, National Kaohsiung University of Applied Sciences, Kaohsiung, Taiwan

[2]Graduate Institute of Education, National Sun Yat-sen University, Kaohsiung, Taiwan

Corresponding Author:
ShuChing Yang, PhD
Graduate Institute of Education
National Sun Yat-sen University
70 Lienhai Road, Kaohsiung 80424, Taiwan, ROC
Kaohsiung,
Taiwan
Phone: 886 752 515 21
Fax: 886 752 558 92
Email: shyang@mail.nsysu.edu.tw

Abstract

Background: College students' health behavior is a topic that deserves attention. Individual factors and eHealth literacy may affect an individual's health behaviors. The integrative model of eHealth use (IMeHU) provides a parsimonious account of the connections among the digital divide, health care disparities, and the unequal distribution and use of communication technologies. However, few studies have explored the associations among individual factors, eHealth literacy, and health behaviors, and IMeHU has not been empirically investigated.

Objective: This study examines the associations among individual factors, eHealth literacy, and health behaviors using IMeHU.

Methods: The Health Behavior Scale is a 12-item instrument developed to measure college students' eating, exercise, and sleep behaviors. The eHealth Literacy Scale is a 12-item instrument designed to measure college students' functional, interactive, and critical eHealth literacy. A nationally representative sample of 525 valid college students in Taiwan was surveyed. A questionnaire was administered to collect background information about participants' health status, degree of health concern, major, and the frequency with which they engaged in health-related discussions. This study used Amos 6.0 to conduct a confirmatory factor analysis to identify the best measurement models for the eHealth Literacy Scale and the Health Behavior Scale. We then conducted a multiple regression analysis to examine the associations among individual factors, eHealth literacy, and health behaviors. Additionally, causal steps approach was used to explore indirect (mediating) effects and Sobel tests were used to test the significance of the mediating effects.

Results: The study found that perceptions of better health status (t_{520}=2.14-6.12, P<.001-.03) and greater concern for health (t_{520}=2.58-6.95, P<.001-.003) influenced college students' development of 3 dimensions of eHealth literacy and adoption of healthy eating, exercise, and sleep behaviors. Moreover, eHealth literacy played an intermediary role in the association between individual factors and health behaviors (Sobel test=2.09-2.72, P<.001-.03). Specifically, higher levels of critical eHealth literacy promoted students' health status and their practice of multiple positive health behaviors, including eating, exercise, and sleep behaviors.

Conclusions: Because this study showed that eHealth literacy mediates the association between individual factors and health behaviors, schools should aim to enhance students' eHealth literacy and promote their health behaviors to help them achieve high levels of critical eHealth literacy. Although some of the study's hypotheses were not supported in this study, the factors that influence health behaviors are complex and interdependent. Therefore, a follow-up study should be conducted to further explore how these factors influence one another.

(J Med Internet Res 2014;16(12):e287) doi:10.2196/jmir.3542

KEYWORDS

demographic; health behavior; mediation; eHealth literacy; quantitative evaluation

XSL•FO
RenderX

Introduction

Background

The health behavior of college students is a topic that is worth exploring. According to Taiwan's Health Promotion Administration of the Ministry of Health and Welfare, National Health Survey statistics showed that the percentage of college students aged 18-24 years who exercised regularly was 60.8% in 2002, 60.6% in 2005, and 55.1% in 2009. Furthermore, the percentage of students who ate breakfast daily was 65.8% in 2002, 62.6% in 2005, and 57.3% in 2009 [1]. Data from a 2010 survey showed a 2-hour difference between the number of hours students spent sleeping on school days (6.4 hours) and weekends (8.5 hours). These results indicate that Taiwanese college students' exercise, eating, and sleep behaviors must be improved [2]. Many of the lifelong habits that jeopardize health are formed during childhood and adolescence [3]. As college students transition from adolescence into adulthood, their health habits may affect their future well-being. During this transition, those who have poor health habits may adopt better habits if they are given sound advice. Therefore, it is necessary and important to examine college students' health behaviors because these behaviors affect students' physical health and lifestyles in adulthood [4]. Health behavior is "any activity undertaken by a person who believes himself to be healthy, for the purpose of preventing disease or detecting it in an asymptomatic stage" [5]. A number of studies on health behavior have focused on eating and exercise [4,6]. Given that sleep is a basic physiological need and an essential element in maintaining physical and mental health, we will examine the health behaviors of college students by measuring their eating, exercise, and sleep behaviors.

Individual factors and eHealth literacy may affect one's health behaviors. According to social cognitive theory [7], each factor possesses a self-regulating system that affects motivation and learner differentiation. Human behavior is influenced and affected by the individual, the behaviors of others, and the environment. This self-regulating system represents a process that is affected by bidirectional and interdependent associations between and among behaviors, environments, and personal experiences. Studies have found that certain factors, such as one's health status, concern for health, and eHealth literacy, may shape an individual's health behaviors [8].

eHealth literacy may mediate the association between demographic factors and health behaviors. The integrative model of eHealth use (IMeHU) suggests that the underlying social structure affects an individual's health status, computer literacy, intrinsic interest in health, and perceived ability to use the Internet for health purposes [9]. The model also proposes macrolevel disparities in the social structures that are connected to health disparities through the microlevel conduits of eHealth literacy, motivation, and ability. That is, individuals with low eHealth literacy have less incentive to use the Internet to access health information and consider themselves to be incapable of using Internet-based health information. The IMeHU provides a parsimonious account of the connections among the digital divide, health care disparities, and the unequal distribution and use of communication technologies. However, few studies have examined the associations among individual factors, eHealth literacy, and health behaviors. Similarly, there is a lack of empirical evidence regarding the IMeHU. This study uses the IMeHU as a framework for examining the associations among individual factors, eHealth literacy, and health behaviors and to further validate the mediating effects of eHealth literacy on health behaviors and individual factors.

Literature Review

Individual Factors Affect Health Behaviors

According to studies of college students in Taiwan, different individual factors affect people's health behaviors [10,11]. These studies found that college students with better self-perceived health status or those enrolled in medical-related fields engaged in healthier behaviors and had more positive health-related attitudes and habits than other students had. For example, medical school students adhere to a more positive, health-promoting lifestyle than do not medical school students [10]. Furthermore, elders who rate their social interactive networks positively demonstrate better health and engage in exercise more consistently [12]. Accordingly, we propose hypothesis H1: based on certain individual factors, the health behaviors of college students can be predicted with relative accuracy.

eHealth Literacy Affects Health Behaviors

In recent years, studies of eHealth literacy have become more prevalent. Some studies focus on defining eHealth literacy [13-15], some design eHealth literacy programs [16,17], and others examine the effect and consequences of eHealth literacy [9]. People with high eHealth literacy are not only more inclined to use the Internet to find answers to health-related questions, but are able to understand the information that they find, evaluate the veracity of the information, discern the quality of different health websites, and use quality information to make informed decisions about health [9]. eHealth literacy affects an individual's health information-seeking behaviors, including the initiative to search for and passively receive messages and then adopt health behaviors based on those messages, which ultimately affects one's health outcomes [18]. Furthermore, those who possess higher levels of eHealth literacy may make healthier decisions, which in turn improve their health outcomes. Researchers have found that the use of health information on the Internet affects personal exercise habits, eating/food consumption habits, and activity habits [19], and that individuals with high eHealth literacy are more likely to permit evaluative procedures for colorectal cancer—a finding that further suggests that those who have better eHealth literacy may adopt more positive health behaviors [20]. Accordingly, we propose hypothesis H2: college students who possess better eHealth literacy engage in more positive health behaviors.

Individual Factors Affect eHealth Literacy

As one's perception of his or her health status improves, his or her health literacy and health knowledge improves [21]. However, some studies show that health status does not affect one's eHealth literacy [17]. Thus, it appears that the effect of health status on one's ability to understand, use, or evaluate

XSL•FO
RenderX

health information is inconclusive and requires further study. Studies on the impact of the frequency of health-related discussions on eHealth literacy show that discussions with parents or peers can promote eHealth literacy [17]. The 2003 US data on the health literacy of American adults found that individuals who often communicate with friends and family and who seek advice from professionals demonstrated higher levels of health literacy [22].

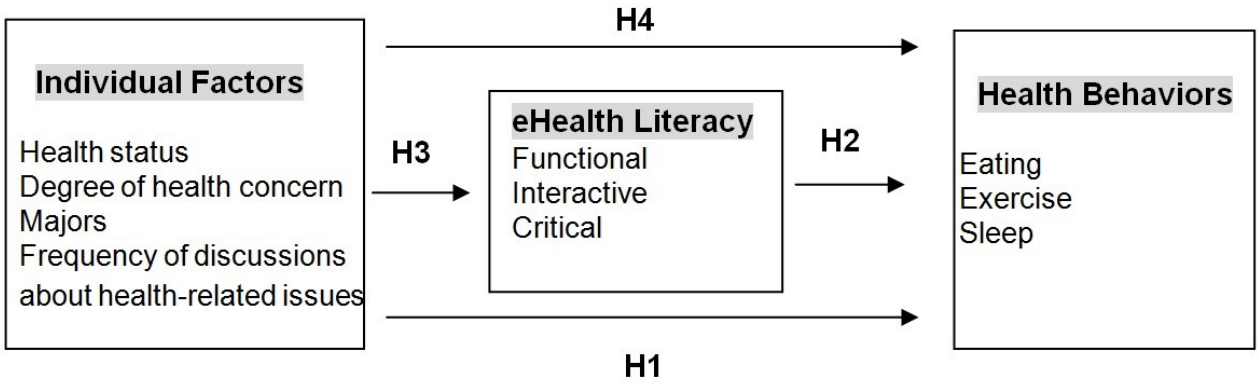
Few studies have examined the effect of students' majors on their eHealth literacy. However, those who are oral health or dental majors tend to have better perceptions of their oral health behaviors than those who are not oral health or dental majors [11]. Similarly, medical school students may have a better cognitive understanding and perception of health information than nonmedical majors. Based on this information, we propose hypothesis H3: among college students, various individual factors (eg, health status, major, degree of health concern, and the frequency with which they engage in health-related discussions) can predict eHealth literacy.

eHealth Literacy Mediates the Association Between Individual Factors and Health Behaviors

Although studies have shown a connection among individual factors, eHealth literacy and health behaviors, no studies have examined the mediating effects. Because eHealth literacy plays an important role in individuals' lives and health behaviors [9], the mediating effects are worthy of further exploration. According to Taiwanese college students, it is difficult for students to evaluate and adopt the suggested behaviors and activities [23]. Even after students critically evaluate the reliability of health information, they often adopt the behavior or activity that has the smallest effect on their health. Thus, hypothesis H4 proposes that eHealth literacy mediates the association between individual factors and health behaviors.

The 4 hypotheses that were developed based on the preceding discussions are shown in Figure 1.

Figure 1. Theoretical framework of the study.



Methods

Participants

According to the Ministry of Education and Statistics Department in Taiwan [24], there were a total of 1,037,632 college students in 2012. Accordingly, a sample of 350 was suggested for this investigation [25]. We determined that 500 students were needed for the official sample and estimated an effective questionnaire return rate of 80%. To make the sample more representative, the stratified cluster sampling method was employed. Using the region on a tiered basis and the class as the sampling unit, we extracted the desired projected number for each region sample based on the proportion of university students in the northern (n=292), central (n=146), southern (n=176), and eastern regions (n=11) of Taiwan. From March to May 2013, we surveyed a nationally representative sample of college students. A total of 625 questionnaires were distributed, including 64 unreturned and 36 incomplete questionnaires. Thus, 525 usable questionnaires were collected, resulting in an effective response rate of 84%. The sample's sociodemographic and health information is displayed in Table 1.

XSL•FO
RenderX

Table 1. Sociodemographic and health information of the sample (N=525).

Variable and group	n (%)
Health status	
Poor (score <4)	21 (4.0)
General (score 5-7)	366 (69.7)
Good (score >8)	138 (26.3)
Degree of health concern	
Extremely unimportant	1 (0.2)
Unimportant	13 (2.5)
Average	215 (40.9)
Important	240 (45.7)
Extremely important	56 (10.7)
Major	
Major in medical field	55 (10.5)
Major in nonmedical field	470 (89.5)
Frequency of discussion about health-related issues	
Never	0 (0.0)
Seldom	118 (22.5)
Sometimes	320 (60.9)
Often	84 (16.0)
Always	3 (0.6)

Instrument

eHealth Literacy Scale

The eHealth Literacy Scale (eHLS) measures a student's ability to seek, find, understand, and evaluate health information from electronic sources and apply this knowledge to address or solve a health problem. The 12-item eHLS, which was developed following a thorough review of the literature [14,26,27], includes the following 3 dimensions: functional, interactive, and critical eHealth literacy. Each dimension is evaluated using 4 items. The functional eHealth literacy dimension evaluates basic reading and writing skills and basic knowledge of health conditions and health systems. Interactive eHealth literacy refers to communicative and social skills that can be used to extract information and derive meaning from different forms of communication and to apply new information to changing circumstances. Critical eHealth literacy refers to advanced cognitive and social skills that can be applied to critically analyze information and to use this information to exert greater control over life events and situations that are related to individual and community-level goals.

An Amos 6.0 confirmatory factor analysis was used to examine the best measurement model. An analysis was conducted with Amos using maximum likelihood estimation. The participants were asked to accurately rate each eHLS item based on their own practices or beliefs using a 5-point Likert scale (1=strongly disagree, 5=strongly agree). We found that the current data adequately fit the eHLS model, which was divided into the following 3 dimensions (12 items total): functional eHealth literacy, interactive eHealth literacy, and critical eHealth literacy. With the use of Amos for the confirmatory factor analysis, a review of the fit indexes revealed a chi-square/df value of 3.02, a goodness of fit index (GFI) value of 0.95, an adjusted GFI (AGFI) value of 0.93, a comparative fit index (CFI) value of 0.95, and a root mean square error of approximation (RMSEA) value of 0.06. Furthermore, the chi-square test was significant (χ^2_{51}=153.8, P<.001).

Health Behavior Scale

The Health Behavior Scale (HBS) was also designed based on the literature [6,28-32] and the reference standard was based on Taiwan Ministry of Healthy Welfare [33] recommendations for individual eating, exercise, and sleep habits. The 3 health behavior dimensions included eating (eg, low-fat dairy foods and low-sugar cereals, vegetable and fruit consumption [more than 5 servings per day]), exercise (eg, exercise at least 3 times per week, monitor pulse while exercising), and sleep behaviors (eg, always have quality sleep, do not fall asleep during the day). Participants responded to the survey questions on a 5-point Likert scale (1=never, 5=always). High scores in the respective dimensions indicated more balanced eating habits, good and regular exercise habits, and good sleep habits and quality.

We found that the current data had a good fit with the model, which was divided into 3 dimensions (12 total items). With the use of Amos to conduct confirmatory factor analysis, a review of the fit indexes revealed a chi-square/df value of 2.74, a GFI value of 0.96, an AGFI value of 0.94, a CFI value of 0.96, and an RMSEA value of 0.06. Furthermore, the chi-square test was

XSL•FO
RenderX

significant (χ^2_{50}=136.8, P<.001). However, a statistically nonsignificant overall chi-square value indicated good fit [34]; these standards reject many models with good fit and suggest other indicators [35]. In this study, the model showed an adequate fit to the data. We considered the 2 models to represent the best-fitting models for the eHLS and HBS structures.

Background Information

Finally, we gathered the respondents' background information, including information about their health status, their major, their degree of health concern, and the frequency with which they engaged in health-related discussions.

The students' health status level (perceived psychological and physical status) was measured by asking them how well they currently felt with respect to their own psychological and physical condition on a scale from 1 (strongly unwell) to 10 (strongly well).

The students' degree of health concern was measured by asking them about their health concerns and willingness to take appropriate action. The degree of importance of participants' health was rated on a scale from 1 (extremely unimportant) to 5 (extremely important). A higher score indicated a higher level of concern regarding one's health.

The major dimension was divided based on whether the participants were majoring in medical fields. In a subsequent analysis, the 2 groups were then transformed into dummy variables. We used the nonmedical group as the reference group.

The frequency with which students engaged in health-related discussions was measured based on participants' responses on a scale from 1 (seldom) to 5 (always).

Data Analysis

This study used Amos 6.0 to perform a confirmatory factor analysis to identify the best measurement models for the eHLS and HBS. We then used multiple regression analysis to examine the associations among individual factors, eHealth literacy, and health behaviors. Additionally, a causal steps approach was used to explore the indirect (mediating) effects [36] and Sobel tests [37] were used to test the significance of the mediating effects.

Results

Multiple Regression Analysis of Individual Factors Predicting eHealth Literacy

Table 2 indicates that nearly all the individual factors positively predicted the 3 dimensions of eHealth literacy, yielding low predictive explanatory powers of 7%-8%. Notably, all the individual factors emerged as significant indicators of the 3 dimensions of eHealth literacy, with the exception of the frequency of health-related discussions.

Table 2. Multiple regression analysis of individual factors predicting eHealth literacy.

Variable	Functional							Interactive							Critical						
	R	ΔR^2	$F_{4,520}$	B	β	t_{520}	P	R	ΔR^2	$F_{4,520}$	B	β	t_{520}	P	R	ΔR^2	$F_{4,520}$	B	β	t_{520}	P
Model summary	.29	.08	11.82					.29	0.8	11.55					.28	.07	10.87				
Health status				.14	.10	2.14	.03				.19	.11	2.34	.02				.28	.11	2.43	.02
Health concern				.37	.14	3.02	.003				.49	.15	3.24	.001				.56	.12	2.58	.01
Majors				.89	.15	3.45	.001				.73	.10	2.31	.02				.92	.09	2.04	.04
Frequency of discussions about health-related issues				.23	.08	1.79	.08				.34	.10	2.16	.03				.64	.13	2.81	.005

Multiple Regression Analysis of Individual Factors and eHealth Literacy Predicting Health Behaviors

Two multiple regression analyses were performed to examine how well individual factors and health literacy predicted health behaviors. Table 3 reveals that the individual factors positively predicted the 3 health behavior dimensions, with a moderate level of predictive explanatory power for eating behaviors (23%) and a low level of predictive explanatory power for exercise (13%) and sleep behaviors (13%).

Table 3 further reveals that both health status and health concern positively predicted the 3 health behavior dimensions; major and the frequency of health-related discussions only demonstrated positive predictive power for eating behaviors. In addition, functional and critical eHealth literacy displayed positive predictive power for eating and exercise behaviors, whereas critical eHealth literacy only positively predicted sleep behaviors.

XSL•FO
RenderX

Table 3. Multiple regression analysis of individual factors and eHealth literacy predicting health behaviors.

Variable	Eating							Exercise							Sleep						
	R	ΔR^2	*F* (*df*)	B	β	*t* (*df*)	*P*	*R*	ΔR^2	*F* (*df*)	B	β	*t* (*df*)	*P*	*R*	ΔR^2	*F* (*df*)	B	β	*t* (*df*)	*P*
Individual factors																					
Model summary	.48	.23	39.37 (4,520)					.37	.13	20.22 (4,520)					.37	.13	20.26 (4,520)				
Health status				.44	.23	5.50 (520)	<.001				.64	.24	5.54 (520)	<.001				.59	.27	6.12 (520)	<.001
Health concern				1.04	.30	6.95 (520)	<.001				.79	.17	3.65 (520)	<.001				.62	.16	3.42 (520)	.001
Majors				.64	.08	2.02 (520)	.04				.71	.06	1.57 (520)	.12				.24	.03	0.64 (520)	.52
Frequency of discussions about health-related issues				.37	.10	3.35 (520)	.02				.30	.06	1.33 (520)	.18				.23	.05	1.21 (520)	.23
eHealth literacy																					
Model summary	.30	.08	17.05 (3,521)					.31	.09	18.03 (3,521)					.26	.06	12.90 (3,521)				
Functional				.21	.16	3.48 (521)	.001				.27	.15	3.23 (521)	.001				.10	.07	1.44 (521)	.15
Interactive				.09	.08	1.64 (521)	.10				.07	.05	0.99 (521)	.32				.07	.06	1.17 (521)	.24
Critical				.11	.14	2.95 (521)	.003				.20	.19	3.98 (521)	<.001				.17	.20	4.03 (521)	<.001

The Mediating Effects of eHealth Literacy on the Association Between Individual Factors and Health Behaviors

Table 4 is the result of a regression analysis of the individual factors and eHealth literacy as predictors of health behaviors. When comparing Tables 2 and 3, the standardized regression coefficients of the independent variables either decreased or were insignificant. Furthermore, a Sobel test of the mediating effect indicated that 6 of 13 paths were significant.

The 2 standardized regression coefficients were multiplied to compute a mediating effect, as presented in Table 5. Critical eHealth literacy mediated the association between health status and eating, exercise and sleep behaviors, showing mediating effect values of .015, .021, and .022, respectively. Functional eHealth literacy mediated the association between health concern and eating and exercise behaviors, with mediating effect values of .022 and .021, respectively. Critical eHealth literacy mediated the association between health concern and eating behaviors, yielding a mediating effect value of .017.

According to Baron and Kenny's approach to statistical mediation analysis [36], it was further determined that critical eHealth literacy mediated the association between health status and health behaviors (see Table 6) and that the mediating effect was .058 units (.015+.021+.022). Functional/critical eHealth literacy mediated the association between health concern and health behaviors, with a mediating effect of .060 units (.022+.021+.017). These results indicate that participants who had better health status and greater concern for their health tended to have better functional and critical health literacy and were, therefore, more inclined to engage in positive health behaviors, especially positive eating behaviors.

XSL•FO
RenderX

Table 4. Regression analysis of individual factors and eHealth literacy predicting health behaviors.

Variable	Eating							Exercise							Sleep						
	R	ΔR^2	$F_{7,517}$	B	β	t_{517}	*P*	*R*	ΔR^2	$F_{7,517}$	B	β	t_{517}	*P*	*R*	ΔR^2	$F_{7,517}$	B	β	t_{517}	*P*
Model summary	.50	.24	25.00					.42	.16	15.65					.41	.15	14.48				
Individual factors																					
Health status				.40	.21	5.02	<.001				.57	.21	4.97	<.001				.54	.25	5.66	<.001
Health concern				.95	.27	6.31	<.001				.62	13	2.90	.004				.52	.13	2.87	.004
Majors				.45	.06	1.43	.15				.38	.04	0.86	.39				.07	.01	0.19	.85
Frequency of discussions about health-related issues				.30	.08	1.87	.06				.15	.03	0.68	.49				.13	.03	0.68	.50
eHealth literacy																					
Functional				.12	.09	2.14	.03				.19	.10	2.31	.02				.04	.02	0.51	.61
Interactive				.03	.02	0.53	.60				.02	.01	0.25	.80				.03	.02	0.41	.68
Critical				.06	.08	1.80	.07				.16	.15	3.23	.001				.13	.15	3.27	.001

Table 5. Sobel test of the mediating effects of eHealth literacy on the association between individual factors and health behaviors.

Pathway	Independent to dependent variable, β		Sobel test	*P*
	Table 1	Table 3		
Health status→functional eHealth literacy→eating	.23	.21	1.96	.05
Majors→functional eHealth literacy→eating	.08	.06	1.32	.19
Health concern→functional eHealth literacy→eating	.30	.27	2.72	<.001
Health status→critical eHealth literacy→eating	.23	.21	2.19	.02
Majors→critical eHealth literacy→eating	.08	.06	1.17	.24
Health concern→critical eHealth literacy→eating	.30	.27	2.39	.01
Frequency of health-related discussions→critical eHealth literacy→eating	.10	.08	1.56	.11
Health status→functional eHealth literacy→exercise	.24	.21	1.95	.05
Health concern→functional eHealth literacy→exercise	.17	.13	2.09	.03
Health status→critical eHealth literacy→exercise	.24	.21	2.18	.02
Health concern→critical eHealth literacy→exercise	.17	.13	1.93	.05
Health status→critical eHealth literacy→sleep	.27	.25	2.23	.02
Health concern→critical eHealth literacy→sleep	.16	.13	0.01	.99

Table 6. Estimation of the causal mediating effects.

Pathway	Estimation of mediated effects
Health status→critical eHealth literacy→eating	.015
Health status→critical eHealth literacy→exercise	.021
Health status→critical eHealth literacy→sleep	.022
Health concern→functional eHealth literacy→eating	.022
Health concern→functional eHealth literacy→eating	.021
Health concern→critical eHealth literacy→eating	.017

XSL•FO
RenderX

Discussion

Overview

This study found that the participants who had better self-perceptions of their health status and stronger concern for their health exhibited better eHealth literacy and had an increased likelihood of adopting healthy eating, exercise, and sleep behaviors. Moreover, eHealth literacy mediated the association between individual factors and health behaviors.

Better Health Status and Greater Health Concern Adopt More Positive Health Behaviors

This study found that the participants who had better perceived health status and greater concern for their health tended to adopt more positive health behaviors. Similarly, college students who majored in medical fields and who engaged in more health-related discussions demonstrated better eating behaviors. Thus, hypothesis 1 was largely supported. Consistent with social cognitive theory [7], an individual's health status and health concern are key factors that prompt him or her to adopt health behaviors [10,12]. It appears that college students who have better perceived health status and greater concern for their own health pay greater attention to their health and are more willing to engage in appropriate health behaviors.

Medical Majors, Better Health Status, and Greater Health Concern Have Better eHealth Literacy Development

The study also revealed that the participants who majored in medical fields, had better perceived health status, and had greater concern for their own health tended to have better eHealth literacy than other students. Additionally, the participants who frequently engaged in health-related discussions had better critical eHealth literacy. Therefore, consistent with previous studies [22,23,38], hypothesis 3 was largely supported: college students who had better perceived health and paid more attention to their health were more likely to seek and evaluate health information.

The framework for health literacy identifies 3 major areas of potential intervention and forms the organizational principles of intervention [8]. The framework also illustrates that educational, health, and cultural and social factors may influence health literacy and may ultimately contribute to health outcomes and costs. Health literacy involves a range of social and individual factors, including cultural and conceptual knowledge, listening, speaking, writing, and reading skills. For example, a previous study found that performance on health literacy tasks was related to education, income, country of birth, age, and race/ethnicity [39]. Specifically, individuals with higher educational attainment and higher income demonstrate higher levels of health literacy [9]. For example, medical school students have higher incomes and more medical knowledge and, therefore, possess greater eHealth literacy than non-medical school students.

No Prediction of Medical Majors and Health-Related Discussions for Exercise and Sleep Behaviors

However, this study also found that the participant's major and the frequency with which he/she engaged in health-related discussions did not predict exercise and sleep behaviors. Because eHealth literacy relates to one's context of relevant medical knowledge when assessing the quality of health information and when making decisions that promote one's health [9], students who major in medical fields are likely to have greater knowledge of medicine. Therefore, medical school students have better health literacy than non-medical school students. The knowledge, attitude, and practice (KAP) model advocates that once an individual receives relevant information, the individual will develop the expected responses, thus triggering behavior that is consistent with one's attitude [40]. However, outcomes and efficacy are not necessarily aligned [7]. In other words, what one knows does not necessarily equate to what one does.

Similarly, the concept of the "KAP gap" states that even if an individual is introduced to new ideas or practices and has a positive attitude toward these new ideas or practices, the individual will not necessarily adopt the behavior [41]. This disconnect results in inconsistency and creates a sense of cognitive dissonance. This problem is especially apparent with respect to new preventive ideas or practices because they often lack positive benefits and advantages. Thus, the effect is more ambiguous and individuals are more prone to experiencing a KAP gap. This tendency may explain why medical majors and those who frequently engage in health discussions do not necessarily implement their knowledge into actions. Future studies are needed to examine the motivation for one's decision to adopt, reject, or discontinue a healthy behavior (ie, what leads to the KAP gap).

Less Influential of Functional and Interactive than Critical eHealth Literacy

Partially consistent with IMeHU [9], we found that critical health literacy played a key role in health status and health behaviors. Functional health literacy positively predicted eating and exercise behaviors, and critical health literacy positively predicted the 3 dimensions of health behaviors. However, interactive health literacy did not contribute to the 3 dimensions of health behavior. Accordingly, hypothesis 2 was partially supported. Individual health information literacy strengthens one's intent to search for and apply health information and influences one's health decision making and engagement in healthy behaviors [18]. Thus, the current study concludes that both functional and critical health literacy and dialog contribute to the adoption of positive health behaviors, especially eating and exercising behaviors.

However, functional eHealth literacy was not found to affect sleep behaviors, and interactive eHealth literacy was not found to affect any of the health behavior dimensions. Therefore, it is inferred that the factors that influence health behaviors are complex. The promotion of individual eHealth literacy affects the retention of eHealth information and subsequently influences future actions [9]. According to involvement theory [42], critical eHealth literacy may motivate individuals to seek and evaluate

XSL•FO
RenderX

the quality of health information. In other words, individuals may attempt to gain as much information as possible and then evaluate and use the information to reach an optimal decision. The processing involved in functional and interactive eHealth literacy does not engage as deeply with an issue as that of critical eHealth literacy. Thus, functional and interactive eHealth literacy are less influential than critical eHealth literacy.

Intermediary Role of eHealth Literacy in the Association Between Individual Factors and Health Behaviors

The findings also revealed that functional and critical eHealth literacy mediated the association between individual factors and health behaviors. Thus, hypothesis 4 was partially supported. We found that functional eHealth literacy only mediated the association between health concern and 2 health behavior dimensions (eating and exercise behaviors). Critical eHealth literacy mediated not only the association between health concern and eating behaviors, but also the association between health status and all 3 health behavior dimensions. Therefore, college students' critical eHealth literacy influenced the association between perceived health status and the 3 health behavior dimensions. These findings are consistent with the belief that critical literacy is crucial to assess the quality of health information because laypersons risk harming themselves in their self-diagnosis and treatment when they lack the required background knowledge to correctly interpret information [43].

Conclusions

This study used the IMeHU to explore the associations among individual factors, eHealth literacy, and health behaviors. We hope that the findings will stimulate further debate about how a health education framework can be translated into practical approaches and will contribute to further refinement of the eHealth literacy concept. The study showed that eHealth literacy played an intermediary role in the association between individual factors and health behaviors. Thus, we suggest that schools strengthen college students' eHealth literacy and promote positive health behaviors based on the current level of eHealth literacy among students [44]. Schools can further use the 6 core skills of eHealth literacy [45] (eg, traditional literacy, health literacy, information literacy, scientific literacy, media literacy, and computer literacy) to develop healthy eating, exercising, and sleeping behavior guidelines and to incorporate these guidelines into health education programs. Moreover, the current findings demonstrated that critical health literacy is a key competence in promoting individual health behaviors. Thus, it is suggested that the development of critical eHealth literacy and the promotion of positive health behaviors among college students require further investigation.

As this study indicated, the IMeHU [9] provides a parsimonious framework that is currently lacking in the extant literature. The examination of the potential effect of eHealth literacy on health behavior presents a unique challenge because it involves a complex interplay of basic literacy skills, the ability to successfully navigate the dominant language framework (English) and culture that is utilized for Web-mediated communication, and sufficient levels of technology adoption and proficiency [46]. Although some research hypotheses failed to gain support in this study, given that the factors that influence health behaviors are complex and interdependent, a future study should be conducted to explore how these factors influence one another. Further research is also needed to refine and verify the parsimonious framework of the IMeHU [9].

Conflicts of Interest

None declared.

References

1. Ministry of Health and Welfare. National Health Survey Statistics. 2009. URL: http://goo.gl/uXxJHU [accessed 2014-02-11] [WebCite Cache ID 6UPqbdzok]
2. Lin L, Chang YY, Ho HS, Yen LL. Sleep problems in adolescents. Formosan J Med 2012;16:72-83.
3. Bandura A. Health promotion from the perspective of social cognitive theory. Psychol and Health 1998;13:623-649.
4. Liou MY, Lu CM. Related factors of health behavior in university students. Chinese J School Health 2006;48:19-37 [FREE Full text]
5. Kasl SV, Cobb S. Health behavior, illness behavior, and sick-role behavior. Archives of Environmental Health: An International Journal 1966 Apr;12(4):531-541. [doi: 10.1080/00039896.1966.10664421]
6. Nomura K, Sato M, Yano E. Developing a measure of communicative and critical health literacy: a pilot study of Japanese office workers. Health Promot Int 2008;23(2):269-274 [FREE Full text]
7. Bandura A. Social cognitive theory: an agentic perspective. Annu Rev Psychol 2001;52:1-26. [doi: 10.1146/annurev.psych.52.1.1] [Medline: 11148297]
8. Nielsen-Bohlman L, Panzer AM, Kindig DA. Health Literacy: A Prescription to End Confusion. Washington, DC: National Academies Press; 2004.
9. Bodie GD, Dutta MJ. Understanding health literacy for strategic health marketing: eHealth literacy, health disparities, and the digital divide. Health Mark Q 2008;25(1-2):175-203. [doi: 10.1080/07359680802126301] [Medline: 18935884]
10. Huang YH, Chiou CJ. Predictors contributing to health-promoting lifestyles among college students in Kaohsiung area. Chinese J Public Health 1997;16(1):24-36.
11. Lu KH, Yen YY, Liu YC, Peng WU, Lin PL, Chen FL, et al. Oral health attitude/ behaviors among dental undergraduate students: an example of a southern medical university. Taiwan J Oral Med Sci 2011;27(1):4-21.

XSL•FO
RenderX

12. Wang KC, Miao, NF, Chen, FY, Chang, YM, Chen, CM. Prevalence of health behaviors and related factors among the elderly of Taiwan. Health Promot Health Educ J 2009;29:73-94.
13. Chan CV, Kaufman DR. A framework for characterizing eHealth literacy demands and barriers. J Med Internet Res 2011 Nov;13(4):e94 [FREE Full text] [doi: 10.2196/jmir.1750] [Medline: 22094891]
14. 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]
15. Stellefson M, Hanik B, Chaney B, Chaney D, Tennant B, Chavarria EA. eHealth literacy among college students: a systematic review with implications for eHealth education. J Med Internet Res 2011 Dec;13(4):e102 [FREE Full text] [doi: 10.2196/jmir.1703] [Medline: 22155629]
16. Hove T, Paek HJ, Isaacson T. Using adolescent eHealth literacy to weigh trust in commercial web sites: the more children know, the tougher they are to persuade. J Advertising Res 2011;51(3):524-537.
17. Paek HJ, Hove T. Social cognitive factors and perceived social influences that improve adolescent eHealth literacy. Health Commun 2012;27(8):727-737. [doi: 10.1080/10410236.2011.616627] [Medline: 22452551]
18. Longo DR. Understanding health information, communication, and information seeking of patients and consumers: a comprehensive and integrated model. Health Expect 2005 Sep;8(3):189-194. [doi: 10.1111/j.1369-7625.2005.00339.x] [Medline: 16098149]
19. Baker L, Wagner TH, Singer S, Bundorf MK. Use of the Internet and e-mail for health care information: results from a national survey. JAMA 2003 May 14;289(18):2400-2406. [doi: 10.1001/jama.289.18.2400] [Medline: 12746364]
20. Mitsutake S, Shibata A, Ishii K, Oka K. Association of eHealth literacy with colorectal cancer knowledge and screening practice among internet users in Japan. J Med Internet Res 2012 Nov;14(6):e153 [FREE Full text] [doi: 10.2196/jmir.1927] [Medline: 23149453]
21. Lee SY, Tsai TI, Tsai YW, Kuo KN. Development and validation of the short-form Mandarin Health Literacy Scale. Taiwan J Pub Health 2012;31(2):184-194.
22. Kutner M, Greenberg E, Baer J. National Center for Education Statistics. 2006. A first look at the literacy of America's adults in the 21st Century (NCES-470) URL: http://nces.ed.gov/NAAL/PDF/2006470.pdf [accessed 2014-02-11] [WebCite Cache ID 6UQA5VOxW]
23. Hsu WC, Chen SF, Ho CJ. Experience of using web health information among college students: an analysis from the health literacy perspective. J Health Promot Health Educ Contents 2011;35:1-22.
24. Department of Statistics of Ministry of Education. Number of students among college and university. 2012. URL: https://stats.moe.gov.tw/files/detail/101/101_student.xls [accessed 2014-02-11] [WebCite Cache ID 6UQA6Mw6V]
25. Creswell JW. Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research. Upper Saddle River, NJ: Pearson/Merrill Prentice Hall; 2008.
26. Chinn D, McCarthy C. All Aspects of Health Literacy Scale (AAHLS): developing a tool to measure functional, communicative and critical health literacy in primary healthcare settings. Patient Educ Couns 2013 Feb;90(2):247-253. [doi: 10.1016/j.pec.2012.10.019] [Medline: 23206659]
27. Ghaddar SF, Valerio MA, Garcia CM, Hansen L. Adolescent health literacy: the importance of credible sources for online health information. J Sch Health 2012 Jan;82(1):28-36. [doi: 10.1111/j.1746-1561.2011.00664.x] [Medline: 22142172]
28. Cho YI, Lee SY, Arozullah AM, Crittenden KS. Effects of health literacy on health status and health service utilization amongst the elderly. Soc Sci Med 2008 Apr;66(8):1809-1816 [FREE Full text] [doi: 10.1016/j.socscimed.2008.01.003] [Medline: 18295949]
29. Kerr MJ, Ritchey DA. Health-promoting lifestyles of English-speaking and Spanish-speaking Mexican-American migrant farm workers. Public Health Nurs 1990 Jun;7(2):80-87. [Medline: 2367289]
30. Arredondo EM, Elder JP, Ayala GX, Campbell N, Baquero B, Duerksen S. Is parenting style related to children's healthy eating and physical activity in Latino families? Health Educ Res 2006 Dec;21(6):862-871 [FREE Full text] [doi: 10.1093/her/cyl110] [Medline: 17032706]
31. Guo X, Warden BA, Paeratakul S, Bray GA. Healthy Eating Index and obesity. Eur J Clin Nutr 2004 Dec;58(12):1580-1586. [doi: 10.1038/sj.ejcn.1601989] [Medline: 15162130]
32. Laberge L, Petit D, Simard C, Vitaro F, Tremblay R, Montplaisir J. Development of sleep patterns in early adolescence. J Sleep Res 2001 Mar 04;10(1):59-67. [doi: 10.1046/j.1365-2869.2001.00242.x]
33. Taiwan Ministry of Health and Welfare. The handbook of healthy lifestyle. 2012. URL: http://bit.ly/1°vPngK [accessed 2014-02-11] [WebCite Cache ID 6UQA9l7fX]
34. Barrett P. Structural equation modelling: adjudging model fit. Pers Indiv Differ 2007;42:815-824 [FREE Full text]
35. Marsh HW, Hau K, Wen Z. In search of golden rules: comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. Struct Equ Modeling 2004;11(3):320-341. [doi: 10.1207/s15328007sem1103_2]
36. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol 1986 Dec;51(6):1173-1182. [Medline: 3806354]
37. Preacher KJ, Leonardelli GJ. Calculation for the Sobel test: an interactive calculation tool for mediation tests. 2003. URL: http://www.quantpsy.org/sobel/sobel.htm [accessed 2014-02-11] [WebCite Cache ID 6UQAaXkHn]

XSL•FO
RenderX

38. Neter E, Brainin E. eHealth literacy: extending the digital divide to the realm of health information. J Med Internet Res 2012;14(1):e19 [FREE Full text] [doi: 10.2196/jmir.1619] [Medline: 22357448]
39. Rudd R, Kirsch I, Yamamoto K. Literacy and Health in America: Policy Information Report. Princeton, NJ: Centre for Global Assessment, Policy Information Centre, Education testing Service; 2004.
40. Allport GW. Attitudes. In: Murchison CM, editor. Handbook of Social Psychology. Worcester, MA: Clark University Press; 1935:798-844.
41. Rogers EM. Diffusion of Innovations. New York: Free Press; 2003.
42. Zaichkowsky J. Measuring the involvement construct. J Consum Res 1985;12:341-352.
43. Kerka S. Health literacy beyond basic skills. 2003. URL: http://eric.ed.gov/?id=ED478948 [accessed 2014-02-11] [WebCite Cache ID 6UQAE86dS]
44. Xie B. Older adults, e-health literacy, and collaborative learning: an experimental study. J Am Soc Inf Sci 2011 Mar 14;62(5):933-946. [doi: 10.1002/asi.21507]
45. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. J Med Internet Res 2006 Jun;8(2):e9 [FREE Full text] [doi: 10.2196/jmir.8.2.e9]
46. Werts N, Hutton-Rogers L. Barriers to achieving E-Health literacy. AJHS 2013;4(3):115-120.

Abbreviations

CFI: comparative fit index
GFI: goodness of fit index
HBS: Health Behavior Scale
IMeHU: integrative model of eHealth use
KAP: knowledge, attitude, and practice
RMSEA: root mean square error of approximation

Edited by G Eysenbach; submitted 15.05.14; peer-reviewed by N Werts, P Boisrond; comments to author 17.07.14; revised version received 30.08.14; accepted 08.11.14; published 12.12.14.

Please cite as:
Hsu W, Chiang C, Yang S
The Effect of Individual Factors on Health Behaviors Among College Students: The Mediating Effects of eHealth Literacy
J Med Internet Res 2014;16(12):e287
URL: http://www.jmir.org/2014/12/e287/
doi:10.2196/jmir.3542
PMID:25499086

©WanChen Hsu, ChiaHsun Chiang, ShuChing Yang. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 12.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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

Movement-Based Interaction Applied to Physical Rehabilitation Therapies

Juan Enrique Garrido Navarro[1], PhD; Victor Manuel Ruiz Penichet[2], PhD; María Dolores Lozano Pérez[2], PhD

[1]Computer Science Research Institute, University of Castilla-La Mancha, Albacete, Spain

[2]Department of Computer Systems, University of Castilla-La Mancha, Albacete, Spain

Corresponding Author:
Juan Enrique Garrido Navarro, PhD
Computer Science Research Institute
University of Castilla-La Mancha
Avenida de España s/n
Albacete, 02071
Spain
Phone: 34 967599200 ext 2656
Fax: 34 967599200
Email: juanenrique.garrido@uclm.es

Abstract

Background: Health care environments are continuously improving conditions, especially regarding the use of current technology. In the field of rehabilitation, the use of video games and related technology has helped to develop new rehabilitation procedures. Patients are able to work on their disabilities through new processes that are more motivating and entertaining. However, these patients are required to leave their home environment to complete their rehabilitation programs.

Objective: The focus of our research interests is on finding a solution to eliminate the need for patients to interrupt their daily routines to attend rehabilitation therapy. We have developed an innovative system that allows patients with a balance disorder to perform a specific rehabilitation exercise at home. Additionally, the system features an assistive tool to complement the work of physiotherapists. Medical staff are thus provided with a system that avoids the need for them to be present during the exercise in specific cases in which patients are under suitable supervision.

Methods: A movement-based interaction device was used to achieve a reliable system for monitoring rehabilitation exercises performed at home. The system accurately utilizes parameters previously defined by the specialist for correct performance of the exercise. Accordingly, the system gives instructions and corrects the patient's actions. The data generated during the session are collected for assessment by the specialist to adapt the difficulty of the exercise to the patient's progress.

Results: The evaluation of the system was conducted by two experts in balance disorder rehabilitation. They were required to verify the effectiveness of the system, and they also facilitated the simulation of real patient behavior. They used the system freely for a period of time and provided interesting and optimistic feedback. First, they evaluated the system as a tool for real-life rehabilitation therapy. Second, their interaction with the system allowed us to obtain important feedback needed to improve the system.

Conclusions: The system improves the rehabilitation conditions of people with balance disorder. The main contribution comes from the fact that it allows patients to carry out the rehabilitation process at home under the supervision of physiotherapists. As a result, patients avoid having to attend medical centers. Additionally, medical staff have access to an assistant, which means their presence is not required in many exercises that involve constant repetition.

(J Med Internet Res 2014;16(12):e281) doi:10.2196/jmir.3154

KEYWORDS

exercise movement techniques; human–computer interaction; interaction devices; movement-based interaction; rehabilitation therapies

XSL•FO
RenderX

Introduction

Background

Doctors, physiotherapists, and health care centers are increasingly incorporating video games and, more specifically, new technologies and methods into their rehabilitation therapies [1-3]. The result is a process that allows patients to recover impaired functions in an entertaining, useful, and effective way. Neurorehabilitation [4] is one of the areas in which this new type of therapy has been used in recent years. In particular, this new rehabilitation procedure has been applied with people with brain damage caused by cardiovascular accident, cerebral infarction, and neurodegenerative diseases such as Parkinson disease, multiple sclerosis, and Alzheimer disease. Although this therapy is currently regarded as unconventional, it has proved highly beneficial when applied in rehabilitation treatment [1,2]. Specifically, the results of a project conducted with patients with multiple sclerosis at Rey Juan Carlos University [5] revealed significant improvement when exercises based on video games were used. Thanks to this new therapy, many patients have been able to fully resume their activities of daily living, such as walking, jumping, or keeping their balance when standing up, that had previously caused them great difficulties.

The use of movement-based interaction is one of the most popular elements of video games that can be applied in health care settings. Specifically, this kind of interaction allows the creation of systems that offer patients the possibility of completing the rehabilitation process at home by following the exercises prescribed by qualified medical staff.

In this paper, we present a system called “BDRehab,” which takes into account how movement-based interaction can improve rehabilitation processes. Specifically, the system focuses on the rehabilitation of patients with neurodegenerative diseases or some form of brain damage (eg, traumatic brain injury, dementia, cerebral palsy). This type of patient is unable to do many of these exercises, such as getting up or sitting on a chair, walking in a straight line, picking up or moving objects. The system focuses on helping patients in the recovery of balance function. These actions, which are easy for people in good health, are complex for people with neurodegenerative and brain diseases. Therefore, rehabilitation and training in these cases represent an essential and important improvement milestone.

The capacity to detect movement enables patient movement to be tracked and analyzed. The results of analysis provide enough information to perform corrective actions. The correction can be achieved by comparing how each rehabilitation exercise is actually carried out with how it should best be done. Accordingly, if the result is correct, then the system communicates this to the patient; if not, the system creates specific advice and clarification in animated form to help the patient complete the rehabilitation process in the correct way. Regarding data collection, the information concerning patient progress in each exercise facilitates the creation of a useful statistical study for medical staff. The aim of this additional functionality is to report on the current state of each patient’s rehabilitation process.

The BDRehab system uses Kinect for Windows [6] to work with movement-based interaction. This device is a major development, as it provides users with the capacity to interact with systems through common and natural gestures. The medical world has found Kinect to be a useful tool with which to experiment to reach solutions and make improvements. A wide range of issues have been treated, from haptic problems [7] to rehabilitation in chronic diseases [8]. The system is generalizable to similar movement-based interaction devices, such as Asus Xtion Pro Live [9] and SoftKinectDepthSense [10].

The rest of the paper is organized as follows. Outstanding related studies are described and compared in the next subsection. Section 2 describes the system developed for rehabilitation at home, as well as its functionality and interface. The outcomes of the evaluation are described in section 3. Section 4 presents conclusions and final remarks.

Related Studies

The emergence of devices offering movement-based interaction in casual environments such as video games has created great interest in their application in medical settings. Specifically, rehabilitation processes incorporating physical exercise comprise one of the main areas in which the research community has found a suitable environment in which to apply these techniques.

The first related study is outstanding. It describes a system, VirtualRehab [11], which involves physical rehabilitation using the movement-based interaction provided by Kinect. The system allows the monitoring and tracking of patients from any location. The main objective of VirtualRehab is to offer patients an enjoyable way of completing complex rehabilitation processes at home. To this end, the equipment required comprises a personal computer, Kinect for Windows software, and a screen. Additionally, VirtualRehab contains a manager that enables medical staff to plan, monitor, and review each patient’s progress. The system focuses on specific pathologies: acquired brain damage, neuromuscular diseases, neurodegenerative diseases, and mobility in older adults. To aid in the treatment of these pathologies, VirtualRehab places patients in a virtual world in which they can work with 9 games to perform specific movements to improve impairments (Figure 1). These games make it possible to work with the affected parts of the body and physical symptoms, particularly in patients with loss of motor ability, movement and posture disorders, and balance and coordination difficulties.

WiiHab is the second system we analyzed. In this case, the technology used is once again based on video games, but uses Nintendo technology through the Wii console. Specifically, WiiHab utilizes the potential of existing Wii games to encourage physical rehabilitation. Many studies [12-14] provide information on positive results with the use of WiiHab. Initially, WiiHab was focused on use of the basic controllers, named “Wii Remote,” to interact with the games. The main games used in physical rehabilitation were those related to sports: tennis, baseball, boxing, golf, and bowling. However, the evolution of the Nintendo console provides the possibility of interaction with new devices, such as the Wii balance board, with which the opportunities of dynamic activities (eg, aerobics or yoga) increase by the use of the feet as an element of interaction.

XSL•FO
RenderX

SeeMe [15] is another major study on rehabilitation, but in this case it corresponds to the assessment and treatment of unilateral spatial neglect. This system represents another method for rehabilitation that creates a virtual reality environment without the requirement of head-mounted displays or specialized equipment. SeeMe is a projected video capture of a virtual story in which patients are "embedded" through their own image. A representation of the patient is generated by capturing the individual through a camera while performing the activities. Specific algorithms are employed for movement and position recognition and analysis. The participants should stand or sit in a specific area while viewing a monitor on which their virtual representation is shown during the exercises inside the virtual world. For example, patients can be embedded in a game in which they are touching virtual exercise balls. Therefore, the participants can see themselves in a virtual story in which they use trunk and limb movements. Additionally, the medical staff is able to change the parameters of the virtual game (based on system indications and staff perceptions) during the rehabilitation procedure to adapt exercises and features according to patient progress and needs. In this way, SeeMe generates a setup whereby patients and medical staff are working together while the patient is at the medical center.

The last system we analyzed is ArmAssist [16], a functional prototype for at-home tele-rehabilitation of poststroke arm impairment. We aimed to provide a portable and easy-to-use device to enable effective rehabilitation (Figure 2). First, the clinical staff needs to decide when the patient is able to use the system independently at home. Then, the device is sent to the patient's home and the rehabilitation process is monitored remotely with appropriate supervision and guidance. The device is integrated into a platform, named "TeleRehab" [17], which supports the phases of therapy planning, training, and assessment. The exercises to be completed with ArmAssist are divided into 2 types: assessment games and training games. The first type (1-2 minutes) is oriented toward working with an objective assessment of the range of movement, vertical force, and the ability to perform specific trajectories. The training games involve more complex activities (5-15 minutes), such as a word completion game. The results are stored and synchronized on a central server to be transmitted to the rehabilitation staff.

The studies mentioned above represent significant advances in rehabilitation processes through the use of new devices, game controllers, and movement-based interactions. However, the system presented in this paper offers new contributions that describe essential differences. In general, BDRehab is part of a complex system, named "Ubi4Health" [18], which provides a comprehensive solution for health care environments by operating from task management to rehabilitation process. Virtual Rehab is the closest work to the BDRehab system. The main difference is that Virtual Rehab uses games instead of real scenarios. Virtual Rehab embeds users in virtual games that help to improve their disabilities. In contrast, the BDRehab system involves patients in the real situation of a common daily activity, such as walking straight. In turn, WiiHab offers a parallel process designed to guarantee a rehabilitation program, but, similarly to Virtual Rehab, it involves patients in different games without engaging the patient in real daily activities. SeeMe represents a highly valid proposal to complete required rehabilitation programs at health care centers. In contrast, BDRehab is able to take the process to the home, which is where patients conduct activities of daily living. Therefore, patients can improve their condition without moving from home and without modifying their habits and comforts. Additionally, SeeMe requires the continuous presence of medical staff during the rehabilitation procedure, whereas BDRehab does not. BDRehab can be considered as an auxiliary application used at home. However, it must always be conducted under the adequate supervision of medical staff who manage the exercises to be completed and analyze the results, modifying parameters if necessary, but without the need to be physically present with the patient. ArmAssist represents a system that involves the use of a device worn by patients on their arms while performing the rehabilitation tasks. In contrast, patients are not required to wear any special equipment while performing their exercises with the BDRehab system. They interact with their own body with the use of Kinect, which captures their movements and postures. Therefore, BDRehab offers comfortable rehabilitation without the need to wear devices.

XSL•FO
RenderX

Figure 1. Example of exercises done in Virtual Rehab [11].

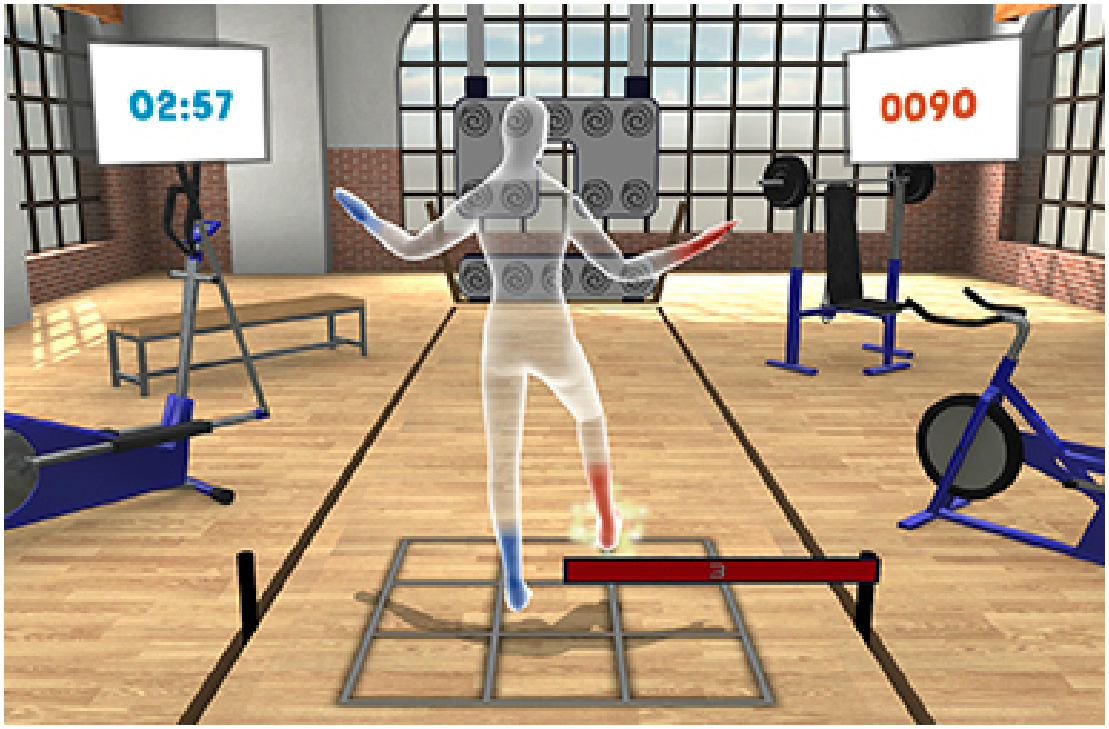


Figure 2. Arm Assist [15].



Methods

Balance Disorder Rehabilitation System

Considering how new technologies, and, more specifically, new interaction modes, can improve the conditions of patients in rehabilitation processes, we have developed a system focused on a specific functional exercise for persons with neurodegenerative diseases. The exercise consists of helping patients to recover from balance disorders by training them to walk straight. Consequently, the system is focused on a very specific objective. Patients are able to complete their therapy at home by performing the exercises defined by the medical staff without having to attend specialized centers. The main achievement is the incorporation of the therapy into the patients' activities of daily living. Over time, the system will become an

XSL•FO
RenderX

element of the patient's home, and rehabilitation will be just another daily activity. However, BDRehab in no way replaces the medical staff responsible for deciding with whom and how each patient has to use the system to perform the rehabilitation tasks. Thus, the decision of when the patient is ready to carry out the process without further supervision is made by the medical staff.

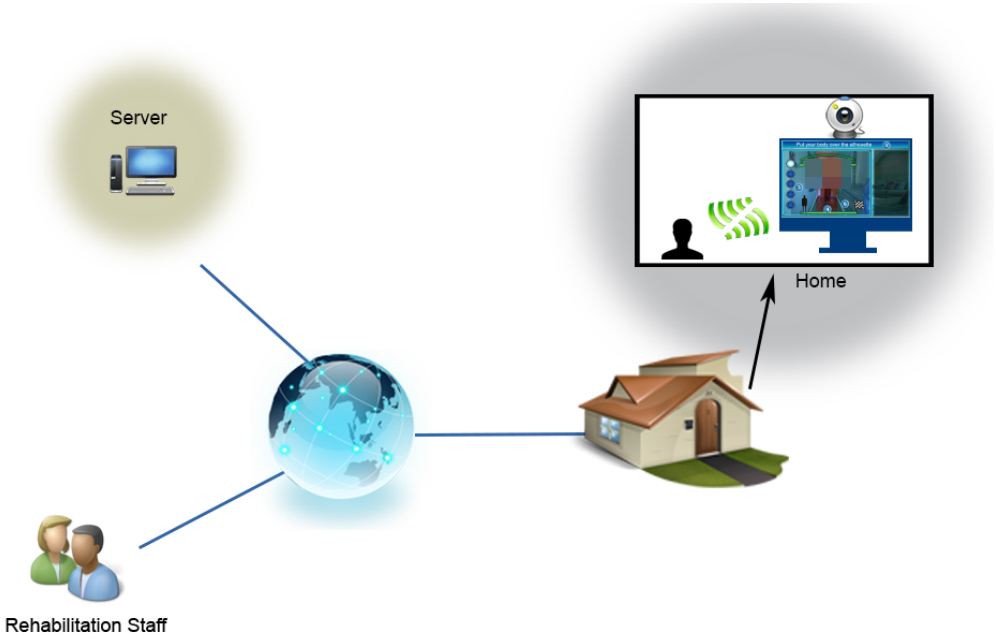
BDRehab can be extended to an important variety of exercises in spite of being focused on just 1 specific exercise. Specifically, the system has been developed to be applied in functional exercises that represent daily life situations. The features of BDRehab support rehabilitation exercises based on physical movements. At present, the movements need to be wide and noticeable, owing to the limitations of the hardware used. However, the evolution in hardware technology has led to increasingly more sensitive devices, meaning that such intense movements will no longer be required for accurate detection.

Based on the information obtained in interviews and meetings with specialists in the field of physiotherapy, one of the keys to success in the recovery of patients with balance disorder, along with many other common problems, is the continuous repetition of the related rehabilitation exercise. This is a major challenge for the patient. These exercises appear simple, but actually entail great complexity. In most cases, the problem stems from the fact that, as one example, the patient is not aware that of losing balance on one side of the body while walking straight. The proposed system monitors patients in real time by allowing them to see themselves in a projected image that provides constant help for performing the series of exercises defined by the medical staff.

The system is a well-defined deployment (see Figure 3) providing the new rehabilitation conditions. In particular, the system deployment is based on 2 connected points: the patient's home computer and the system server. The first point represents the place where each Kinect device operates. To set up the system at the patient's home, only a personal computer connected to a Kinect device and to a screen is needed. The personal computer requires a Web connection with a minimum bandwidth of 64 kbps to access the server, Windows 7 as the operating system, a dual core processor, and 2 GB of random access memory. In terms of visualization, the screen shows the system interface, which is the part of the system responsible for guiding the patient during the rehabilitation process. The server supplies and receives information to and from the software system running in the personal computer. Specifically, the information received is the usage statistics for each performed exercise to help medical staff analyze the evolution of the rehabilitation processes: attempts, errors, and the time needed to perform each session. The information supplied serves the purposes of parameter setting to customize the exercises for the patient.

The next subsections describe 3 essential elements of the system: (1) the exercises performed by the patients; (2) the system functionality with description of the fundamentals of the development and how it works; and (3) the interface that is needed to offer a system that helps patients and physiotherapists in the best possible way.

Figure 3. System Deployment.



Balance Exercises

A number of real rehabilitation exercises comprise the exercise to be performed by the patient. These are then checked by the system as an adaptation, including the most important aspects. The exercises to be conducted using the system were selected according to recommendations in the specialized literature [19,20] as well as the knowledge and experience of the physiotherapists involved. The experience of these professionals allowed the exercises to be adapted to the needs of rehabilitation patients.

One of the exercises consists of walking along a straight line without leaning sideways and going backward in the same way.

XSL•FO
RenderX

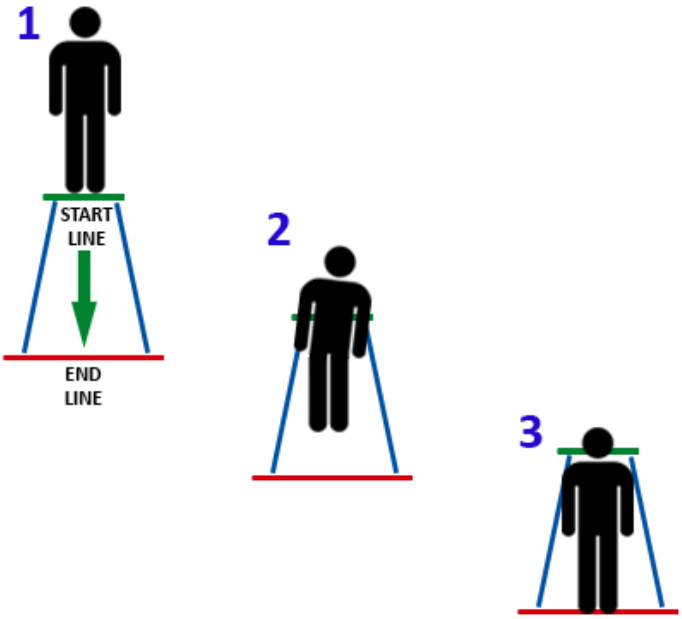
In general, guide marks help the patients line up the area within which they have to walk. These marks are an essential element in reporting a deviation of which the patient is not aware.

The other exercise consists of improving balance through ensuring an upright posture of for a number of seconds without any swinging movements. In this case, the patient has to be in a specific posture: standing up, arms close to the body without lifting, back straight, and feet aligned with the shoulders.

The BDRehab system combines the previous exercises to generate a more complex exercise based on the first one. However, it monitors patients to ensure they maintain an upright posture while walking (as required by the second exercise) and also when still; in other words, when they start and finish the exercise.

Figure 4 shows the steps that comprise the complete BDRehab rehabilitation exercise. First, the patient has to stay still in an upright posture. Second, the patient has to walk along a 5-meter straight line without losing balance or deviating from the line. The last step consists of stopping at the finish line and maintaining an upright posture as at the start.

Figure 4. Steps to be performed during the BDRehab exercise: (1) to start in an upright posture; (2) to go along five-meter long straight line; and (3) to stop at the finish line and maintaining the upright posture.



System Functionality

The objective of the system is to check the posture of patients uninterruptedly during the exercise to guide them to perform the exercise described in the previous subsection. The system considers the exercise in 3 separate parts that involve different checks related to the postures and movements of the patient. To this end, the system maintains a variable named “step” to know which part of the exercise is performed and then the conditions to check. The system checks if the patient is at the initial position when the variable has the value 1. For value 2, the system checks if the patient is in the middle of the circuit and is balancing the body. For value 3, the system checks the patient’s final position and posture. In this way, the value of the variable increases from 1 to 3 after each step is checked. This monitoring process is carried out for each repeated exercise.

The monitoring process, which is recorded within a specific framework, comprises 2 parts. First, the system has to identify the posture of the patient. Posture identification is carried out by using the software development kit (SDK) provided by Microsoft [21]. This process is a key element, so an explanation of how the system performs the posture identification is essential. The SDK recognizes all 20 joints of the body that Kinect is able to identify. These points are supplied to the system through a code element that represents the skeleton and provides access to them through a set of 3 values referring to the position of the point in space (*x*-, *y*-, and *z*-axes). Once the system obtains this code element, the next step is to check if the patient is in the required posture, depending on the step of the exercise. To this end, the system considers that there are pairs of body parts that may be aligned with the vertical-horizontal coordinates. For example, the points related to the shoulders have to be at the same height during the exercise; otherwise, the patient is inclined to the side related to the shoulder that is in a lower position. However, the comparison allows a low error rate, which is checked through the difference of such coordinates within an error range.

An essential feature of this framework is that it works independently of patient height. The main reason for this is that the comparison of coordinates is relative to each patient. The framework checks if a pair of body points is aligned in the *x*- or *y*-axis, which means the coordinates are the same in one of these axes or inside the error range allowed. In other words, the difference between the *x*/*y*-coordinates of 2 points, which may be aligned in the *x*/*y*-axis, must be zero or inside the error range, regardless of the height of the patient. Figure 5 shows an example in which the right feet and shoulders of 2 patients of different height are analyzed to establish if these body parts

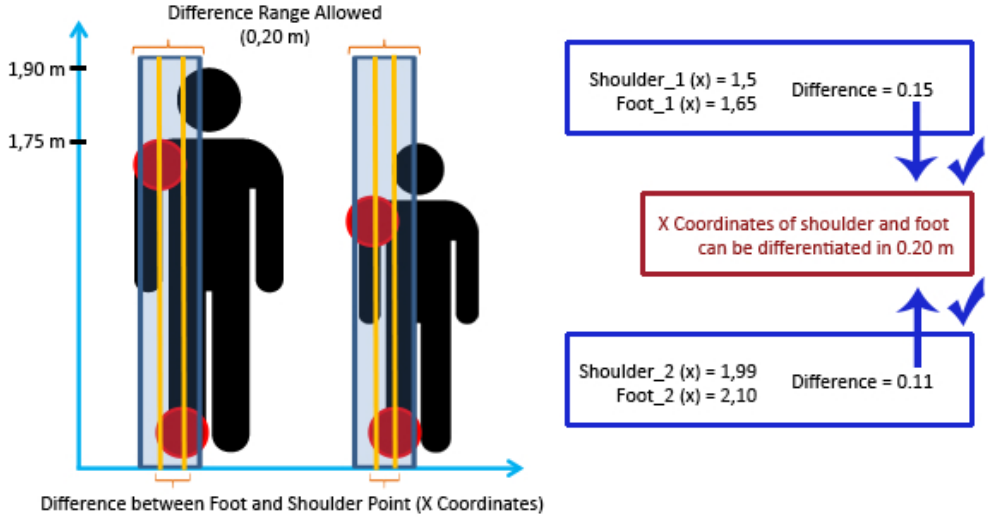
XSL•FO
RenderX

with the same *x*-coordinate are within an error range. The figure contains 2 marks for each patient to check if the posture is correct: (1) the range within which the difference of the *x*-coordinates of the related body points should be and (2) the difference itself. In this case, as the difference is within the range, the posture is correct for the 2 patients. The figure shows 2 patients for whom the framework has analyzed a specific posture, regardless of the patients' difference in height. The important idea is that the 2 analyzed points should be of a similar value along the *x*-axis.

The use of difference of coordinates allows the framework to be independent of patient height. This feature avoids the need to calibrate the system for each user, a task that previously had to be carried out by every patient before starting the rehabilitation process. Therefore, patients are able to perform the rehabilitation session without any other complications.

The accuracy of the monitoring process is controlled through the management of a set of thresholds. The described framework and the whole system work under the control of each threshold, which can be customized by the physiotherapists. Increasing or decreasing each threshold makes the framework more or less strict. Specifically, the modification of the thresholds involves 2 factors: consideration of (1) wider or narrower error ranges as well as (2) the whole body of the patient or just some specific parts of it when checking the posture. This method modifies the precision of the system to detect postures and, consequently, the precision with which the patients have to perform the rehabilitation session. Therefore, the patients' performance evolves favorably from the beginning to the end of the process. Initially, when patients find the exercise and the interaction more difficult, parameters are set at the minimum threshold; as they learn and overcome their limitations, the thresholds can be modified and increased accordingly. In this way, the thresholds are used to modify BDRehab based on the limitations on the patient's movements. In technical terms, Kinect provides a degree of accuracy that is considered to be sufficient, natural, and adequate for the rehabilitation process carried out with the BDRehab system.

Figure 5. Example of the algorithm to check correct postures.



System Interface

Overview

A good user interface is an essential element for patient acceptance of the system. Consequently, a rehabilitation system needs to be easy to understand and needs to guide patients adequately during the exercise under the physiotherapist's supervision. As a result, patients will be more likely to include the system as a rehabilitation tool in their activities of daily living. In this sense, the system can guide each patient step by step, just as physiotherapists would do. Additionally, the interface can incorporate enough information to complete the specific exercise and also to help guide and correct users at any point of the process, giving them the corresponding feedback.

The system interface is based on the interaction provided by Kinect. The interface follows the generic usability guidelines for any application [22], as well as those defined by the Kinect device [23]. The generic guidelines refer to the visual aspect of the system, which must allow the patient to avoid uncomfortable working positions. Furthermore, the Kinect guidelines point the way to creating helpful interfaces controlled by natural movements, such as hand and head movements.

The type of user is a key component in analyzing the needs and appearance of the interface. Some patients may understand current technology, whereas others may know nothing about it. In this sense, analysis has shown that it is fundamental that the interface have a well-defined set of warnings, notifications, and signals. At the same time, the system can show the information on the screen together with appropriate metaphors to complete the indications. Therefore, the system has to constantly let patients know exactly what to do and how to do it, particularly because the process of the exercise is controlled by patient interaction.

Figures 6 and 7 show screenshots of the system in which it is possible to observe the information given to patients during the rehabilitation process. Each element of the interface is outlined in the paragraphs below.

XSL•FO
RenderX

Current Task of the Exercise to be Performed

The patient needs to know what to do at all times. The system analyzes the development of the exercise by studying the distance covered by the patient. First, the system checks that the user is at the starting point. If the patient is not at the required point, the system indicates that the patient needs to be there. Then, the system indicates that the next step is to walk until the end of the rehabilitation route. Finally, the system determines the end of the exercise when the patient reaches the end of the virtual line that has been followed.

Balance Level

This mark is a key piece of information. A set of scales that inclines to the side on which the patient is leaning while walking shows the patient's balance. This is an effective way of showing the correctness of the position. If the user inclines the body excessively, the balance is inclined and the color of the affected side changes to red.

Attempts

The system records the attempts the user has made to complete the exercise, based on the physiotherapist's indications.

Level of Completion of the Complete Exercise

Patients need to know which stage of the exercise they are at. This signal marks a man in a line showing the progress made. At the left is the starting point, and at the right is the finishing point.

Margins of Movement

These elements are 2 squares that diffuse the image to limit the action range of the patient during the walk. The margins widen as the patient nears the finishing point to generate a correct sense of depth. If the patient touches any margin, it will light up in red to indicate what is happening.

Walking Lines

Walking elements create margins to show the walking space available to perform the exercise. If the user touches any of the lines, the system lights up the corresponding margin in red to show what is happening.

Patient's Silhouette

This element helps patients locate their bodies when adopting the initial posture. Specifically, the silhouette uses a section in red to represent the possible position of the patient.

Animation Area

This area contains the animations with a guide highlighted in red. The animation shows the patient the next action to be performed. This area includes information regarding the postures and movements that need to be corrected.

Misaligned Bones Corrector

This element indicates badly positioned joints during the current exercise using a semitransparent red surface. In this way, the corrector helps the patient understand the errors they make.

Corrective Arrows

The corrective arrows indicate the side toward which the patient's posture must be corrected during the exercise. Figure 7 shows how posture must be corrected to the appropriate side.

Conclusions

This information appears when the patient finishes each attempt. This element gives a list of what needs to be corrected in the next step. If the attempt is performed correctly, the system "congratulates" the patient.

It is important to highlight that, to avoid causing the patient unnecessary stress, the time recording is not shown to the patient. However, BDRehab internally collects the duration of each exercise attempt, which is seen only by the medical staff. They can use the time recording as additional information to modify the accuracy of the monitoring process through the set of thresholds and also to better assess the evolution of the patients.

The possibility of using the patient's voice to control the functionality of the system completes the interface. This capacity is provided by the Kinect device, which can interact with users through their voice by offering a microphone component and related software with voice recognition capability. This creates a better environment for users with arm mobility problems, in which they can operate the interface using oral instructions. Primarily, it allows users to manage the system by using preferences or simple actions such as indicating that they are ready to start.

XSL•FO
RenderX

Figure 6. System screenshot which shows user instructions in order to complete the rehabilitation exercise. The interface elements in the screenshot are: (a) current task, (b) balance level, (c) attempts, (d) level of completion, (e) margins of movement, (f) walking lines and (g) patient's silhouette.



Figure 7. System screenshot which shows an overview of a complete rehabilitation exercise. The interface elements in the screenshot are: (h) animation area, (i) misaligned bones corrector and (j) corrective arrows.



Results

The evaluation of the system has been performed following a user-centered approach. The process comprised 2 iterations. The outcomes of both evaluations were utilized to develop the system. Taking this into account, the version of the system described in this paper reflects most of the improvements that the evaluation process provided as an iterative process.

Five simulated patients took part in each iteration of the evaluation process. According to Nielsen and Landauer [24], 5

XSL•FO
RenderX

participants are sufficient to find 85% of errors and/or problems. The group comprised 2 teenagers (14 and 16 years old), 1 adult (31 years old), and 2 older adults (70 and 75 years old). The environment consisted of a 7×4-meter room in which a 24-inch screen connected to a Kinect device was installed at a height of 1 meter.

Two therapists were present during the whole evaluation process. They were responsible for ensuring that the process was carried out correctly. The active involvement of the therapists was a key factor in the success of the evaluations, because, unlike the present authors, they are experts in balance disorder rehabilitation and were better able to manage the process. The main objective was to know the effectiveness of the system. Therapists are fully familiar with the behavior patterns and limitations of the patients and were thus able to guide the participants to imitate the movements of a real patient with great precision.

The evaluation was recorded using a video camera to collect each participant's reaction when using the system. In addition, the image displayed on the screen was recorded to show the relationship between the participant's reactions and what the system was displaying on the screen.

The evaluations were performed in 2 iterations. The first one was performed without any prior explanation to the patients about the operation mode of the BDRehab system. In this way, we aimed to obtain information about the intuitiveness of the interface elements. The participants had to detect, identify, and interpret each element and functionality to complete one attempt of the rehabilitation exercise. Next, a complete explanation about BDRehab was given to the participants, and they then attempted the exercise 4 times. Finally, the participants were asked about the intuitiveness and simplicity of the system and what they thought could be eliminated.

The outcomes provided interesting feedback that provided a basis on which to make improvements to the system following the established user-centered design approach. We collected participants' impressions, which are graphically depicted in Figure 8. In general, most participants found the system easy to work with. There was only one case in which a participant could not identify an element of the interface, which was the balance level. All other participants were able to identify the elements totally or partially. However, some initial interface elements (44%, 4 of 9) presented difficulties, because the partial identification was greater than or equal to the total identification. Specifically, 4 elements of the interface were involved in most problems: balance level, animation area, patient silhouette, and corrective arrows.

Only one color was used for the balance level, and many participants felt the color should change when their balance moved toward one side, indicating a motion error. The animation area confused older people when the examples showed an image in a sideways position. The same participants had problems following the patient silhouette because some points vibrated, causing them to be easily distracted. According to these findings, we and the therapists improved the following 4 aspects of the system. (1) The balance level was changed to green if the posture is correct and red if the posture moves to an incorrect side. (2) The animation area was changed, using new animations that better represented the postures to be repeated. (3) The code of the system was analyzed and corrected to avoid any vibration during the exercises. (4) The corrective arrows were modified to make them curved to represent the curvature of the back.

During the first iteration of the evaluation, the 2 therapists conducted a thorough analysis of the system. As a result, they considered the system a useful tool for the patients during their rehabilitation programs. The system avoids the need for patients to frequently attend medical centers and allows them to complete rehabilitation at home if their conditions are suitable. Despite positive impressions, the therapists suggested a set of improvements to be made, taking into account how patients usually work. First, they found it necessary to consider the possibility of working with patients with disabilities in other parts of the body. Specifically, the system was highly restrictive and failed if patients did not move an arm during the exercise. The solution found was to allow therapists to customize the body parts to be checked during the exercise. In addition, the therapists emphasized the need to improve the method of indicating the direction of the user during the exercise. The main reason was that many patients have difficulty in maintaining the same direction from the starting point to the end. Hence, we included 2 elements in the interface: the margins of movements and the walking lines (Figure 6). An additional consideration was the need to include more information about the exercise based on errors made. In this way, a more detailed report about the errors made during the exercise was included in the conclusions element.

The second iteration consisted of using a new version of the BDRehab system with many differences from the initial one. The objective was to know whether the redesigned interface provided a suitable set of elements and consequently whether the results of the first iteration had been interpreted correctly. The same participants attempted and successfully completed the rehabilitation exercise 4 times. Two additional improvements were also established. The older participants required audio instructions, as they experienced problems when listening to the instructions while checking their own exercises on the screen. This modification was introduced as an essential element because it includes additional information that helps patients complete the rehabilitation exercise without cluttering the user interface. The therapists found the system appropriate for application in a rehabilitation program but considered that some patients prefer not to see their own body displayed on the screen, so they thought the system would benefit from a virtual environment to eliminate real images. This concern is an interesting possibility for future research.

Finally, we tested the usability of the latest version of the system, in terms of user satisfaction, by using the System Usability Scale (SUS) [25]. The results of this test are shown in Figure 9. SUS scores have a range of 0 to 100, and we achieved a score of 98, which means that the usability of the system is reasonably high.

XSL•FO
RenderX

Figure 8. Identification of the interface elements.

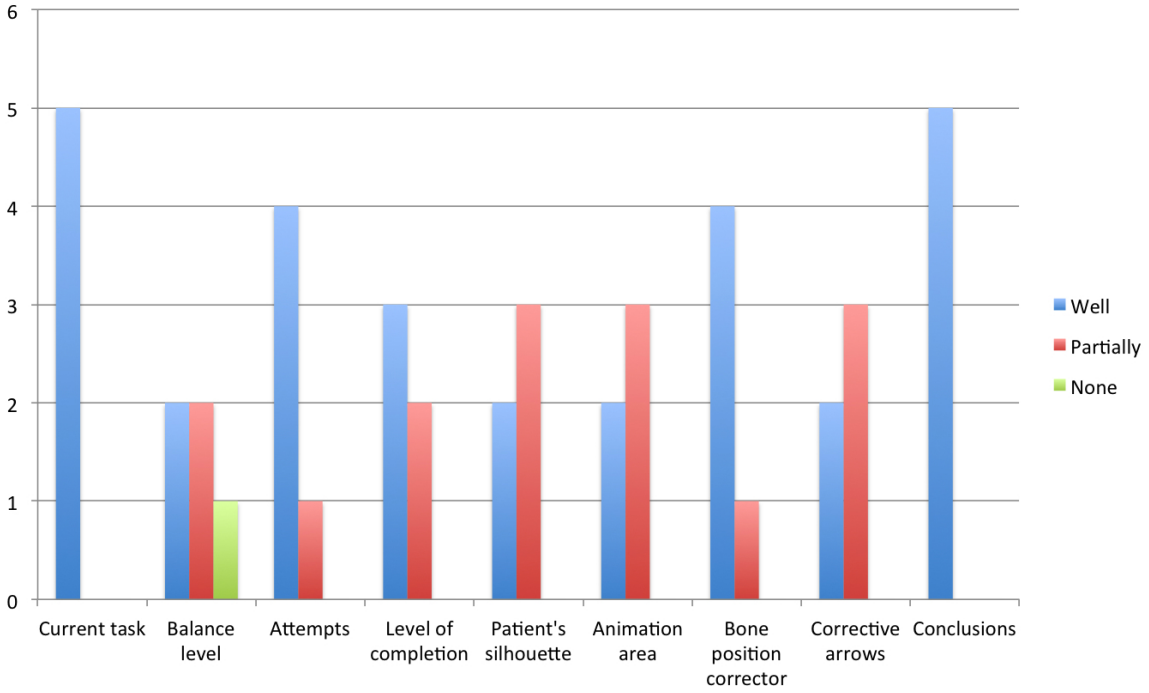
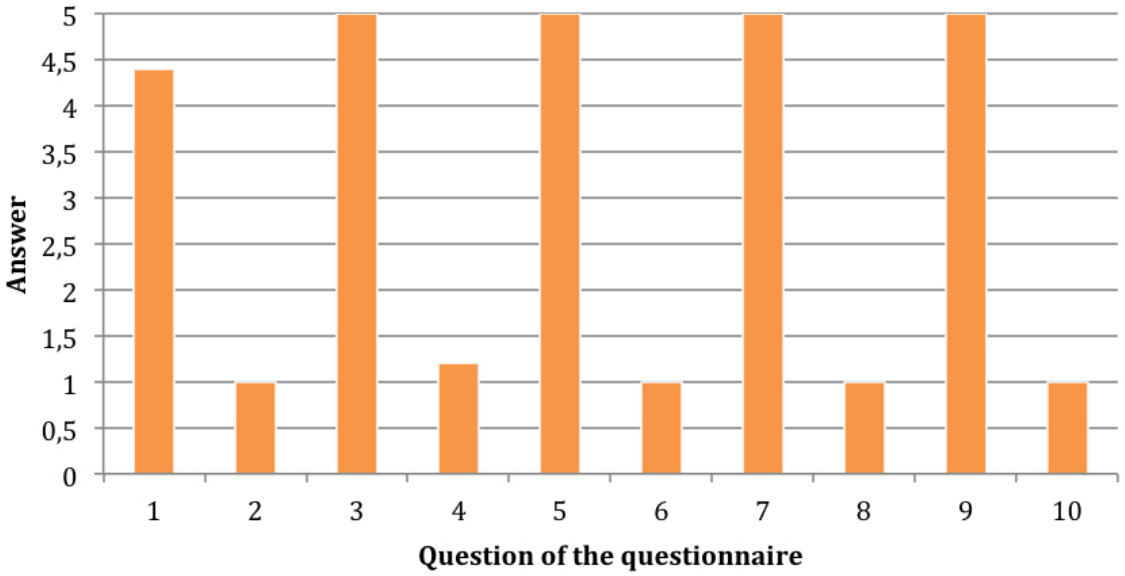


Figure 9. System Usability Scale (SUS) Results.



Discussion

In this paper, we present a system designed to improve rehabilitation conditions of people with balance disorders. In particular, the system serves as a rehabilitation tool for patients who need training to walk straight. The main contribution comes from the fact that the system allows patients to perform the rehabilitation process at home under the supervision of physiotherapists.

The proposal presents significant benefits from two points of view. The first is that it takes into account the perspective of the medical staff, who find the system to be a useful time-saving tool. The time gained increases the capacity of medical staff to attend to patients more intensively, defining specific parameters for the exercise for each patient. The system partially replaces the presence of therapists during the performance of the exercise. Therapists are then able to focus on the results and the way to improve each rehabilitation process. The rehabilitation exercise is usually highly repetitive. This means that the constant presence of the physiotherapist is not required, because instructions are provided by the system. These instructions provide patients with enough information for them to know what they are doing incorrectly and how to correct the action. In this way, the system contains different ways to alert the patients without overwhelming them. The easiest instructions, and generally the most common ones, appear on the screen,

XSL•FO
RenderX

adding color to the part of the exercise that is being wrongly completed. For example, the system uses an arrow to show the way to correct balance (see Figure 7). Other indications are more intrusive if the problem to be corrected could affect the patient's health. For example, patients may fall if they do not correct a certain action. Although rehabilitation requires constant monitoring, traditionally with the physical presence of a therapist, this system constantly monitors the repetitive exercises, replicating the full-time attention usually provided by a therapist.

On the one hand, one might think that, without the presence of medical staff patients miss out, to some degree, on the personal treatment that can often be of great comfort. On the other hand, the patients are able to make progress at home with their family or people close to them. Despite the fact that medical staff is not physically present, they receive updated information about the progress of each patient. BDRehab collects information about patients' progress and generates statistical data on how patients are performing in each attempt during the rehabilitation process, allowing patients to control their progress and analyze the results in depth. Regarding how medical staff can assess the progress of the rehabilitation sessions, the system stores the data produced during the rehabilitation session at the patient's home. The therapists can then analyze these data and modify the parameters of the exercises accordingly to adapt the exercises to the progress of the patient.

The second point of view to be considered is that of the patient. The system avoids the need for patients to attend medical or specialty centers for rehabilitation. In this way, the patients can remain at home, a familiar and more comfortable environment, which is certainly of physiological benefit. Additionally, patients can conduct the rehabilitation process in an independent and less invasive way without the need to depend on others more than is strictly necessary. Therefore, they can organize their progress at their own leisure, based on the steps marked out by medical staff. These 2 benefits represent a significant development in rehabilitation, especially taking into account that some patients have to continue the rehabilitation process for the rest of their lives.

This proposal builds upon movement-based interaction through Kinect devices and a framework that analyzes patients' movements. The framework studies how each patient performs the exercise and, depending on this performance, creates appropriate instructions to provide feedback to the patient to correct and guide them. The analysis consists of checking the posture and movements of the patient in the 3 steps into which the rehabilitation exercise is divided. The first and last steps help check if the patient has the correct posture at the beginning and end of the exercise: standing up, arms close to the body, back straight, and feet aligned with the shoulders. The intermediate steps consist of analyzing how the patient walks from the beginning to end. In this sense, the system searches for straight movements. Feedback instructions are fundamental to help the patient during the exercise. The data gathered during the performance of the exercise (eg, time, number of errors) are stored so the therapist can supervise the evolution of the patient's progress and adapt the exercises accordingly.

The system has been evaluated by 2 therapists with outstanding experience in balance disorder rehabilitation and by 5 people representing real patients. The outcomes of the evaluation were positive, indicating that the system could be applied in a real rehabilitation program as a useful aid for patients to follow the therapy at home and also as a suitable auxiliary tool for therapists to supervise the evolution of patients' progress without the need to be physically present all the time. Furthermore, the therapists involved in the research provided us with interesting ideas for improving the system, which were crucial to achieving the final version described in this paper.

Acknowledgments

This work was partially funded by project TIN2011-27767-C02-01 from the Spanish Ministry of Economy and Competitiveness and by project TSI-100101-2013-147 from the Spanish Ministry of Industry, Energy and Tourism.

Conflicts of Interest

None declared.

References

1. Parry IS, Bagley A, Kawada J, Sen S, Greenhalgh DG, Palmieri TL. Commercially available interactive video games in burn rehabilitation: therapeutic potential. Burns 2012;38(4):493-500. [doi: 10.1016/j.burns.2012.02.010] [Medline: 22385641]
2. Smith ST, Schoene D. The use of exercise-based videogames for training and rehabilitation of physical function in older adults: current practice and guidelines for future research. Aging Health 2012;8(3):243-252. [doi: 10.2217/ahe.12.30]
3. Marston HR, Smith ST. Interactive Videogame Technologies to Support Independence in the Elderly: A Narrative Review. Games for Health Journal 2012;1(2):139-152. [doi: 10.1089/g4h.2011.0008]
4. van Paasschen J, Clare L, Yuen KS, Woods RT, Evans SJ, Parkinson CH, et al. Cognitive rehabilitation changes memory-related brain activity in people with Alzheimer disease. Neurorehabil Neural Repair 2013;27(5):448-459. [doi: 10.1177/1545968312471902] [Medline: 23369983]
5. Luna-Oliva L, Ortiz-Gutiérrez R, Cano-de la Cuerda R, Martínez Piédrola R, Alguacil-Diego IM, Sánchez-Camarero C. Evaluation of the use of a virtual reality video-game system as a supplement for rehabilitation of children with cerebral

XSL•FO
RenderX

palsy. In: Pons JL, Torricelli D, Pajaro M, editors. Converging Clinical and Engineering Research on Neurorehabilitation (Biosystems & Biorobotics series Volume 1). Berlin, Germany: Springer Berlin Heidelberg; 2013:873-877.
6. Microsoft. Kinect for Windows. SDK 2. URL: http://www.microsoft.com/en-us/kinectforwindows/ [accessed 2013-11-28] [WebCite Cache ID 6LTLtDJo5]
7. Shrewsbury BT. Providing haptic feedback using the Kinect. In: ASSETS '11: Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility; 2011 Oct 24-26; Dundee, UK. New York, NY: ACM; 2011 Presented at: 13th International ACM SIGACCESS Conference on Computers and Accessibility; 2011 Oct 24-26; Dundee, UK p. 321-322. [doi: 10.1145/2049536.2049628]
8. Schönauer C, Pintaric T, Kaufmann H, Jansen-Kosterink S, Vollenbroek-Hutten M. Chronic pain rehabilitation with a serious game using multimodal input. In: Proceedings of the 2011 International Conference on Virtual Rehabilitation. New York, NY: IEEE; 2011 Presented at: 2011 International Conference on Virtual Rehabilitation; 2011 Jun 27-29; Zurich, Switzerland URL: http://publik.tuwien.ac.at/files/PubDat_204331.pdf
9. ASUS. Xtion PRO LIVE. Source: ASUS URL: http://www.asus.com/Multimedia/Xtion_PRO_LIVE/ [accessed 2013-11-28] [WebCite Cache ID 6LTM7G5RO]
10. SoftKinetic. Source: SoftKinetic. DepthSense cameras URL: http://www.softkinetic.com/en-us/products/depthsensecameras.aspx [accessed 2013-11-28] [WebCite Cache ID 6LTMO47lo]
11. Virtualware Group. Source: Virtualware Group. VirtualRehab URL: http://virtualrehab.info/product/ [accessed 2013-11-30] [WebCite Cache ID 6LWIRck1u]
12. Herz NB, Mehta SH, Sethi KD, Jackson P, Hall P, Morgan JC. Nintendo Wii rehabilitation ("Wii-hab") provides benefits in Parkinson's disease. Parkinsonism Relat Disord 2013;19(11):1039-1042. [doi: 10.1016/j.parkreldis.2013.07.014] [Medline: 23968649]
13. Fung V, So K, Park E, Ho A, Shaffer J, Chan E, et al. The utility of a video game system in rehabilitation of burn and nonburn patients: a survey among occupational therapy and physiotherapy practitioners. J Burn Care Res 2010;31(5):768-775. [doi: 10.1097/BCR.0b013e3181eed23c] [Medline: 20628305]
14. Shih CH, Shih CT, Chu CL. Assisting people with multiple disabilities actively correct abnormal standing posture with a Nintendo Wii balance board through controlling environmental stimulation. Res Dev Disabil 2010;31(4):936-942. [doi: 10.1016/j.ridd.2010.03.004] [Medline: 20381997]
15. Sugarman H, Weisel-Eichler A, Burstin A, Brown R. Use of novel virtual reality system for the assessment and treatment of unilateral spatial neglect: a feasibility study. New York, NY: IEEE; 2011 Presented at: Proceedings of the 2011 International Conference on Virtual Rehabilitation; 2011; New York p. 27-29. [doi: 10.1109/ICVR.2011.5971859]
16. Perry JC, Zabaleta H, Belloso A, Rodríguez-de-Pablo C, Cavallaro FI, Keller T. ArmAssist: development of a functional prototype for at-home telerehabilitation of post-stroke arm impairment. In: Proceedings of the 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob). New York, NY: IEEE Robotics and Automation Society; 2012:1561-1566.
17. Perry JC, Ruiz-Ruano JA, Keller T. Telerehabilitation: toward a cost-efficient platform for post-stroke neurorehabilitation. IEEE Int Conf Rehabil Robot 2011;2011:5975413. [doi: 10.1109/ICORR.2011.5975413] [Medline: 22275616]
18. Garrido JE, Penichet VMR, Lozano MD. Integration of collaborative features in ubiquitous and context-aware systems using distributed user interfaces. In: Distributed User Interfaces: Usability and Collaboration (Human-Computer Interaction Series). London, UK: Springer London; 2013:69-83.
19. Stokes M. Physiotherapy to Neurorehabilitation. Madrid, Spain: Elsevier; 2006.
20. Levitt S. Cerebral Treatment and Motor Disease. Barcelona, Spain: Panamericana; 1996.
21. Shotton J, Fitzgibbon A, Cook M, Sharp T, Finocchio M, Moore R, et al. Real-time human pose recognition in parts from single depth images. In: Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition Jun 20-25; Colorado Springs, CO. Washington, DC: IEEE Computer Society; 2011:1297-1304.
22. Constantine L, Lockwood L. Software for Use: A Practical Guide to the Models and Methods of Usage-Centered Design. Reading, MA: Addison Wesley; 1999.
23. Microsoft. Kinect for Windows Human Interface Guidelines. Version 1. URL: http://msdn.microsoft.com/en-us/library/jj663791.aspx [accessed 2014-10-31] [WebCite Cache ID 6TjqlKmJa]
24. Nielsen J, Landauer TK. A mathematical model of the finding of usability problems. In: Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems Apr 24-29; Amsterdam, The Netherlands. New York, NY: ACM; 1993:206-213.
25. Brooke J. SUS: a "quick and dirty" usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland IL, editors. Usability Evaluation in Industry. London, UK: Taylor & Francis; 1996:189-194.

Abbreviations

SDK: software development kit
SUS: System Usability Scale

XSL•FO
RenderX

Edited by G Eysenbach; submitted 30.11.13; peer-reviewed by H Fardoun, C Vincent, WC Su; comments to author 11.02.14; revised version received 07.04.14; accepted 17.10.14; published 09.12.14.

Please cite as:
Garrido Navarro JE, Ruiz Penichet VM, Lozano Pérez MD
Movement-Based Interaction Applied to Physical Rehabilitation Therapies
J Med Internet Res 2014;16(12):e281
URL: http://www.jmir.org/2014/12/e281/
doi:10.2196/jmir.3154
PMID:25491148

©Juan Enrique Garrido Navarro, Victor Manuel Ruiz Penichet, María Dolores Lozano Pérez. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 09.12.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), 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

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

XSL•FO
RenderX