Journal of Medical Internet Research

Journal Impact Factor (JIF) (2023): 5.8
Volume 17 (2015), Issue 7 ISSN 1438-8871 Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Original Papers

Online Alcohol Assessment and Feedback for Hazardous and Harmful Drinkers: Findings From the AMADEUS-2 Randomized Controlled Trial of Routine Practice in Swedish Universities (e170)
Preben Bendtsen, Marcus Bendtsen, Nadine Karlsson, Ian White, Jim McCambridge. 4

A Web-Based, Social Networking Physical Activity Intervention for Insufficiently Active Adults Delivered via Facebook App: Randomized Controlled Trial (e174)
Carol Maher, Monika Ferguson, Corneel Vandelanotte, Ron Plotnikoff, Ilse De Bourdeaudhuij, Samantha Thomas, Karen Nelson-Field, Tim Olds. 33

Engagement and Nonusage Attrition With a Free Physical Activity Promotion Program: The Case of 10,000 Steps Australia (e176)
Diana Guertler, Corneel Vandelanotte, Morwenna Kirwan, Mitch Duncan. 47

Impact of Baseline Assessment Modality on Enrollment and Retention in a Facebook Smoking Cessation Study (e179)
Andrea Villanti, Megan Jacobs, Grace Zawistowski, Jody Brookover, Cassandra Stanton, Amanda Graham. 61

Effect of a Web-Based Behavior Change Program on Weight Loss and Cardiovascular Risk Factors in Overweight and Obese Adults at High Risk of Developing Cardiovascular Disease: Randomized Controlled Trial (e177)
Sinead Watson, Jayne Woodside, Lisa Ware, Steven Hunter, Alanna McGrath, Christopher Cardwell, Katherine Appleton, Ian Young, Michelle McKinley. 66

The Mobile Insulin Titration Intervention (MITI) for Insulin Adjustment in an Urban, Low-Income Population: Randomized Controlled Trial (e180)
Natalie Levy, Victoria Moynihan, Annielyn Nilo, Karyn Singer, Lidia Bernik, Mary-Ann Etiebet, Yixin Fang, James Cho, Sundar Natarajan. . . . 85

It's LiFe! Mobile and Web-Based Monitoring and Feedback Tool Embedded in Primary Care Increases Physical Activity: A Cluster Randomized Controlled Trial (e184)
Sanne van der Weegen, Renée Verwey, Marieke Spreeuwenberg, Huibert Tange, Trudy van der Weijden, Luc de Witte. 109

A Web-Based Adolescent Positive Psychology Program in Schools: Randomized Controlled Trial (e187)
Rowan Burckhardt, Vijaya Manicavasagar, Philip Batterham, Leonie Miller, Elizabeth Talbot, Alistair Lum. 123

Medium-Term Effectiveness of a Comprehensive Internet-Based and Patient-Specific Telerehabilitation Program With Text Messaging Support for Cardiac Patients: Randomized Controlled Trial (e185)
Ines Frederix, Dominique Hansen, Karin Coninx, Pieter Vandervoort, Dominique Vandijck, Niel Hens, Emeline Van Craenenbroeck, Niels Van Driessche, Paul Dendale. 134

XSL•FO
RenderX

Overcoming Clinical Inertia: A Randomized Clinical Trial of a Telehealth Remote Monitoring Intervention Using Paired Glucose Testing in Adults With Type 2 Diabetes (e178)
Deborah Greenwood, Shelley Blozis, Heather Young, Thomas Nesbitt, Charlene Quinn. 165

Allowing Physicians to Choose the Value of Compensation for Participation in a Web-Based Survey: Randomized Controlled Trial (e189)
Alison Turnbull, Cristi O'Connor, Bryan Lau, Scott Halpern, Dale Needham. 195

Information and Communication Technology Use Among Low-Income Pregnant and Postpartum Women by Race and Ethnicity: A Cross-Sectional Study (e163)
Nymisha Chilukuri, Meredith West, Janice Henderson, Shari Lawson, Robert Ehsanipoor, Kathleen Costigan, Sarah Polk, Wendy Bennett. . .
2 0 5

Prevalence and Frequency of mHealth and eHealth Use Among US and UK Smokers and Differences by Motivation to Quit (e164)
Belinda Borrelli, Yvonne Bartlett, Erin Tooley, Christopher Armitage, Alison Wearden. 215

Mobile Phone Sensor Correlates of Depressive Symptom Severity in Daily-Life Behavior: An Exploratory Study (e175)
Sohrab Saeb, Mi Zhang, Christopher Karr, Stephen Schueller, Marya Corden, Konrad Kording, David Mohr. 243

The Web-Based Osteoarthritis Management Resource My Joint Pain Improves Quality of Care: A Quasi-Experimental Study (e167)
Hema Umapathy, Kim Bennell, Chris Dickson, Fiona Dobson, Marlene Fransen, Graeme Jones, David Hunter. 254

Why Breast Cancer Risk by the Numbers Is Not Enough: Evaluation of a Decision Aid in Multi-Ethnic, Low-Numerate Women (e165)
Rita Kukafka, Haeseung Yi, Tong Xiao, Parijatham Thomas, Alejandra Aguirre, Cindy Smalletz, Raven David, Katherine Crew. 267

Improving Pediatric Basic Life Support Performance Through Blended Learning With Web-Based Virtual Patients: Randomized Controlled Trial (e162)
Ronny Lehmann, Christiane Thiessen, Barbara Frick, Hans Bosse, Christoph Nikendei, Georg Hoffmann, Burkhard Tönshoff, Sören Huwendiek.
2 7 8

Scientific Versus Experiential Evidence: Discourse Analysis of the Chronic Cerebrospinal Venous Insufficiency Debate in a Multiple Sclerosis Forum (e159)
Janka Koschack, Lara Weibezahl, Tim Friede, Wolfgang Himmel, Philip Makedonski, Jens Grabowski. 290

Twitter Social Media is an Effective Tool for Breast Cancer Patient Education and Support: Patient-Reported Outcomes by Survey (e188)
Deanna Attai, Michael Cowher, Mohammed Al-Hamadani, Jody Schoger, Alicia Staley, Jeffrey Landercasper. 302

Epidemic Wave Dynamics Attributable to Urban Community Structure: A Theoretical Characterization of Disease Transmission in a Large Network (e169)
Anne Hoen, Thomas Hladish, Rosalind Eggo, Michael Lenczner, John Brownstein, Lauren Meyers. 320

Characterizing the Processes for Navigating Internet Health Information Using Real-Time Observations: A Mixed-Methods Approach (e173)
Susan Perez, Debora Paterniti, Machelle Wilson, Robert Bell, Man Chan, Chloe Villareal, Hien Nguyen, Richard Kravitz. 332

Secure Cloud-Based Solutions for Different eHealth Services in Spanish Rural Health Centers (e157)
Isabel de la Torre- Díez, Miguel Lopez-Coronado, Begonya Garcia-Zapirain Soto, Amaia Mendez-Zorrilla. 343

XSL•FO
RenderX

A Cost-Effectiveness Analysis of Blended Versus Face-to-Face Delivery of Evidence-Based Medicine to Medical Students (e182)
Stephen Maloney, Peter Nicklen, George Rivers, Jonathan Foo, Ying Ooi, Scott Reeves, Kieran Walsh, Dragan Ilic. 357

Viewpoints

How to Increase Reach and Adherence of Web-Based Interventions: A Design Research Viewpoint (e172)
Geke Ludden, Thomas van Rompay, Saskia Kelders, Julia van Gemert-Pijnen. 19

Trials of Intervention Principles: Evaluation Methods for Evolving Behavioral Intervention Technologies (e166)
David Mohr, Stephen Schueller, William Riley, C Brown, Pim Cuijpers, Naihua Duan, Mary Kwasny, Colleen Stiles-Shields, Ken Cheung. . . .
2 3 0

A Privacy Preservation Model for Health-Related Social Networking Sites (e168)
Jingquan Li. 310

Reviews

Evaluation of Internet-Based Interventions on Waist Circumference Reduction: A Meta-Analysis (e181)
Dong-Chul Seo, Jingjing Niu. 96

Adverse Drug Reaction Identification and Extraction in Social Media: A Scoping Review (e171)
Jérémy Lardon, Redhouane Abdellaoui, Florelle Bellet, Hadyl Asfari, Julien Souvignet, Nathalie Texier, Marie-Christine Jaulent, Marie-Noëlle Beyens, Anita Burgun, Cédric Bousquet. 149

Online Recruitment Methods for Web-Based and Mobile Health Studies: A Review of the Literature (e183)
Taylor Lane, Julie Armin, Judith Gordon. 182

Letter to the Editor

Photoaging Mobile Apps: A Novel Opportunity for Smoking Cessation? (e186)
Titus Brinker, Werner Seeger. 367

XSL•FO
RenderX

Original Paper

Online Alcohol Assessment and Feedback for Hazardous and Harmful Drinkers: Findings From the AMADEUS-2 Randomized Controlled Trial of Routine Practice in Swedish Universities

Preben Bendtsen[1*], PhD; Marcus Bendtsen[2*], MSc; Nadine Karlsson[3], PhD; Ian R White[4], PhD; Jim McCambridge[5*], PhD

[1]Medical Faculty, Department of Medical Specialist and Department of Medicine and Health Sciences, Linköping University, Motala, Linköping, Sweden

[2]Technical Faculty, Department of Computer and Information Science, Linköping University, Linköping, Sweden

[3]Medical Faculty, Department of Medicine and Health Sciences, Linköping University, Linköping, Sweden

[4]MRC Biostatistics Unit, Cambridge Institute of Public Health, Cambridge, United Kingdom

[5]Faculty of Science, Department of Health Sciences, Universiyt of York, Heslington, United Kingdom

[*]these authors contributed equally

Corresponding Author:
Preben Bendtsen, PhD
Medical Faculty
Department of Medical Specialist and Department of Medicine and Health Sciences
Linköping University, Motala
Campus US
Linköping, Se-581 83
Sweden
Phone: 46 702324615
Email: preben.bendtsen@liu.se

Related Article:

This is a corrected version. See correction statement: http://www.jmir.org/2016/5/e118

Abstract

Background: Previous research on the effectiveness of online alcohol interventions for college students has shown mixed results. Small benefits have been found in some studies and because online interventions are inexpensive and possible to implement on a large scale, there is a need for further study.

Objective: This study evaluated the effectiveness of national provision of a brief online alcohol intervention for students in Sweden.

Methods: Risky drinkers at 9 colleges and universities in Sweden were invited by mail and identified using a single screening question. These students (N=1605) gave consent and were randomized into a 2-arm parallel group randomized controlled trial consisting of immediate or delayed access to a fully automated online assessment and intervention with personalized feedback.

Results: After 2 months, there was no strong evidence of effectiveness with no statistically significant differences in the planned analyses, although there were some indication of possible benefit in sensitivity analyses suggesting an intervention effect of a 10% reduction (95% CI –30% to 10%) in total weekly alcohol consumption. Also, differences in effect sizes between universities were seen with participants from a major university (n=365) reducing their weekly alcohol consumption by 14% (95% CI –23% to –4%). However, lower recruitment than planned and differential attrition in the intervention and control group (49% vs 68%) complicated interpretation of the outcome data.

Conclusions: Any effects of current national provision are likely to be small and further research and development work is needed to enhance effectiveness.

Trial Registration: International Standard Randomized Controlled Trial Number (ISRCTN): 02335307; http://www.isrctn.com/ISRCTN02335307 (Archived by WebCite at http://www.webcitation.org/6ZdPUh0R4).

XSL•FO
RenderX

(J Med Internet Res 2015;17(7):e170) doi:10.2196/jmir.4020

KEYWORDS

alcohol drinking; behavior therapy; students; Internet; electronic mail, feedback

Introduction

College and university students in Sweden, as in other parts of the world, drink alcohol heavily [1-2]. Because alcohol is responsible for substantial adverse health consequences and social problems [3-4], there is a need for effective interventions. Swedish universities offer preventive services aiming at reducing drinking among students through 26 local student health care centers nationally. Human resources (ie, staff numbers) are not sufficient to offer face-to-face brief interventions to all risky (hazardous) or problem (harmful) drinkers, although such interventions have been shown to have small beneficial effects [5-6]. Even if offered, take-up can be expected to be low and the Internet offers promise for wider-reaching and cost-effective interventions [6,7].

Existing evidence of the effectiveness of online interventions with students is mixed [6-8] partly because of unresolved methodological challenges including attrition prevention and assessment reactivity [9]. In a narrative review in 2008 on online alcohol interventions targeting students, Elliott et al [10] found an effect that was better than no intervention and equivalent with other alcohol interventions. In another review in 2009, Carey et al [6] reviewed the effect of online alcohol intervention including 26 studies targeting students and found an overall short-term reduction in alcohol consumption with weighted mean effect sizes for various alcohol measures from approximately 0.13 to 0.29 that decreased over time. A significant variability in efficacy was seen due to a heterogeneity of content, tailoring, and method of access to the intervention (ie, logging on to a website on home computer or performing the intervention in an office-based setting) [6]. In a later meta-analysis in 2009 of 43 online interventions to student populations, the interventions were found to reduce both quantity and frequency measures of consumption with small effect sizes (0.09 to 0.28) over short (5 weeks or less) and long-term intervals (6 weeks or more) [11]. In a review in 2011 including 19 randomized controlled trial (RCT) studies where the student population represented the largest proportions of participants, a significant reduction in weekly alcohol consumption and binge drinking were found in student populations, but the findings were tentative because of methodological weaknesses in the studies [8]. In the most recent review in 2014 including 23 studies in which most were performed among student populations, online alcohol interventions were found effective in reducing consumption up to a 12-month period with a mean difference in consumption of approximately 1.5 to 2 standard drinks of alcohol [7].

After a large pilot study that successfully addressed study design issues [12-14], the AMADEUS-1 trial in Sweden [15-16] targeted both risky and nonrisky drinkers in a nontreatment-seeking student population and showed small but beneficial effects of assessment in comparison with a no-contact control group, with little additional impact of feedback [16]. AMADEUS-1 used an unconventional trial design with students unaware they were participating in a trial [15]. We preferred not to alter the intervention content of the national system based on this single evaluation study with an unconventional trial design. These considerations led us to design the subsequent study, AMADEUS-2, as a conventional 2-arm RCT design targeting hazardous and harmful drinkers only, using a single screening question and no baseline assessment to minimize assessment reactivity [4,17]. AMADEUS-2 [18] aims to provide a further evaluation of the national system for online alcohol intervention used in routine practice among the student health care centers in Sweden, specifically among the key target population of hazardous and harmful drinkers.

Methods

Study Design and Hypothesis

The study was a 2-arm parallel group RCT in which routine provision of single-session online alcohol assessment and feedback intervention (Group 1) was compared with nonintervention (Group 2) by delaying access to the intervention for 2 months until research follow-up was completed. Sweden is the first country to implement a national system of proactive alcohol intervention for students via student health care services. However, the timing of intervention delivery varies across Sweden and we took advantage of this lack of standardization of timing to implement random allocation at the individual level in this effectiveness evaluation study. The primary hypothesis was that the intervention group would reduce their total weekly alcohol consumption compared to the control group after 2 months. Ethical approval for the study was granted by the Regional Ethical Committee in Östergötland, Sweden (no: 2013/46-31).

Study Procedures Including Recruitment and Randomization

The study was undertaken in 9 of 26 student health care centers in Sweden, each providing services to one university or college. These centers were selected on the basis that they had not previously been involved in RCTs in our research program. All students in their 2nd, 4th, or 6th terms (n=54,507) at the 9 colleges/universities were sent an email in March 2013 inviting them to answer a single question about their drinking. If eligible for trial participation, they were provided with information permitting fully informed consent, making this study unlike AMADEUS-1. The number of students invited varied across the colleges/universities according to their size from 831 in the smallest college (Gävle) to 13,102 in the largest university (Lund). The single screening question used was the third item of the Alcohol Use Disorders Identification Test (AUDIT) questionnaire on the frequency of heavy episodic drinking (HED) [19], used here with a 3-month timeframe. Single alcohol screening questions have been validated as identifying hazardous and harmful drinkers in different settings [20-21] and this type

XSL•FO
RenderX

of drinking is particularly important in this population [22]. Students who were drinking 5 standard drinks (12 grams of alcohol in Sweden) or more for men or 4 standard drinks or more for women twice a month or more often were deemed eligible for trial participation. This approach was previously used by Walters and colleagues [23] who were similarly concerned to avoid reactivity to screening. The key underlying problem to be avoided was that assessment and intervention effects could interact to bias estimates of the effects of behavioral interventions in trials [15,16,24].

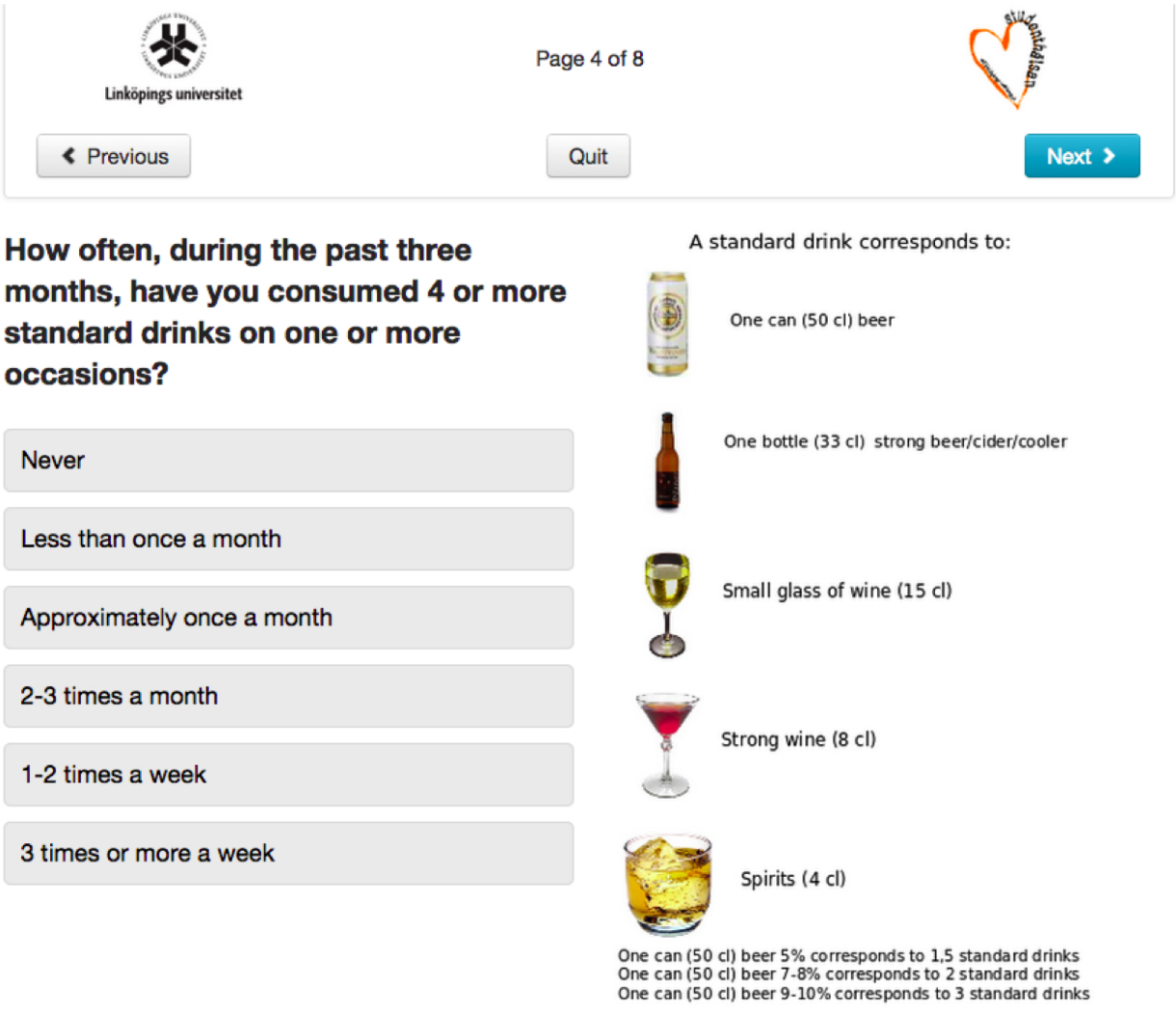
The initial email routinely sent from the participating student health care center was altered to invite study participation. Eligible consenting students were immediately randomized to intervention or control conditions. Randomization was done using Java's built in random number generator (java.util.Random); thus, randomization was fully computerized, did not employ any strata or blocks, and all subsequent study processes were fully automated (programmed by MB). Unlike the AMADEUS-1 trial [14], there was no blinding in this study. The former group gained immediate access to the intervention and the latter group were informed that they would be able to access the intervention in 2 months' time. Two months later, both groups were sent an identical email by the researchers. This email contained an invitation to participate in the follow-up survey, which included the same questions and type of feedback used in the baseline intervention received by the intervention group 2 months earlier. There were a total of 4 reminders (making 5 opportunities to respond in all), initially at weekly intervals, then at shorter intervals, with the final email making clear that this was the last opportunity to respond and allowing 2 days to do so. There were no incentives used to encourage study participation or retention.

Intervention Content

Immediately after randomization, the intervention group only were asked to complete an assessment. Four questions were asked about sex, age, domestic situation, and faculty of study. Alcohol consumption was calculated as total number of standard drinks for each of the 7 days in a typical week during the last 3 months (this intervention component was also later used as the primary outcome in the trial and note was unavailable for the control group to minimize reactivity); other questions explored frequency of HED, the largest amount of alcohol intake in standard drinks on a single occasion during the last 3 months, negative experiences perceived to be related to alcohol, and motivation to reduce alcohol consumption (Figures 1-3). Participants then received feedback consisting of 3 statements summarizing their weekly consumption, their frequency of HED, and their highest blood alcohol concentration during the last 4 weeks, comparing drinking patterns against the safe drinking limits established by the Swedish Institute for Public Health [25]. Also, a graphic illustration of their level of risk was given using green, yellow, and red colors to indicate risk status. After this followed comprehensive normative feedback with information describing participants' alcohol use compared to their peers in Swedish universities (adjusted for sex and age group) and, if applicable, personalized advice on reducing unhealthy levels or patterns of consumption. The feedback was shown on the screen and could also be printed out by the student. A PDF version of the feedback was also emailed to the students immediately after closing this page. A demonstration version in English of the assessment and feedback intervention can be viewed online [26].

XSL•FO
RenderX

Figure 1. Screenshot of AMADEUS: assessment of heavy episodic drinking.



XSL•FO
RenderX

Figure 2. Screenshot of AMADEUS: assessment of weekly drinking.

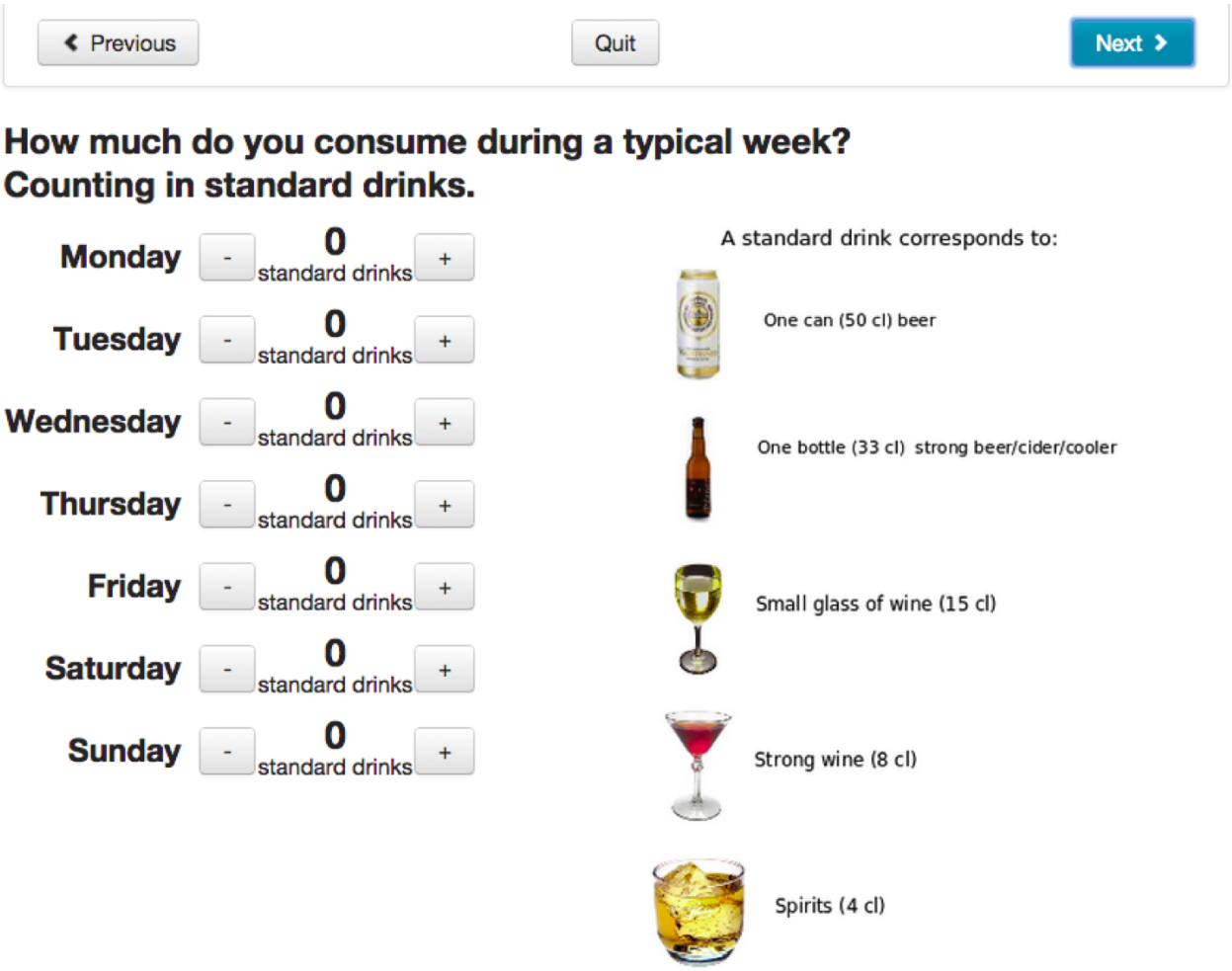
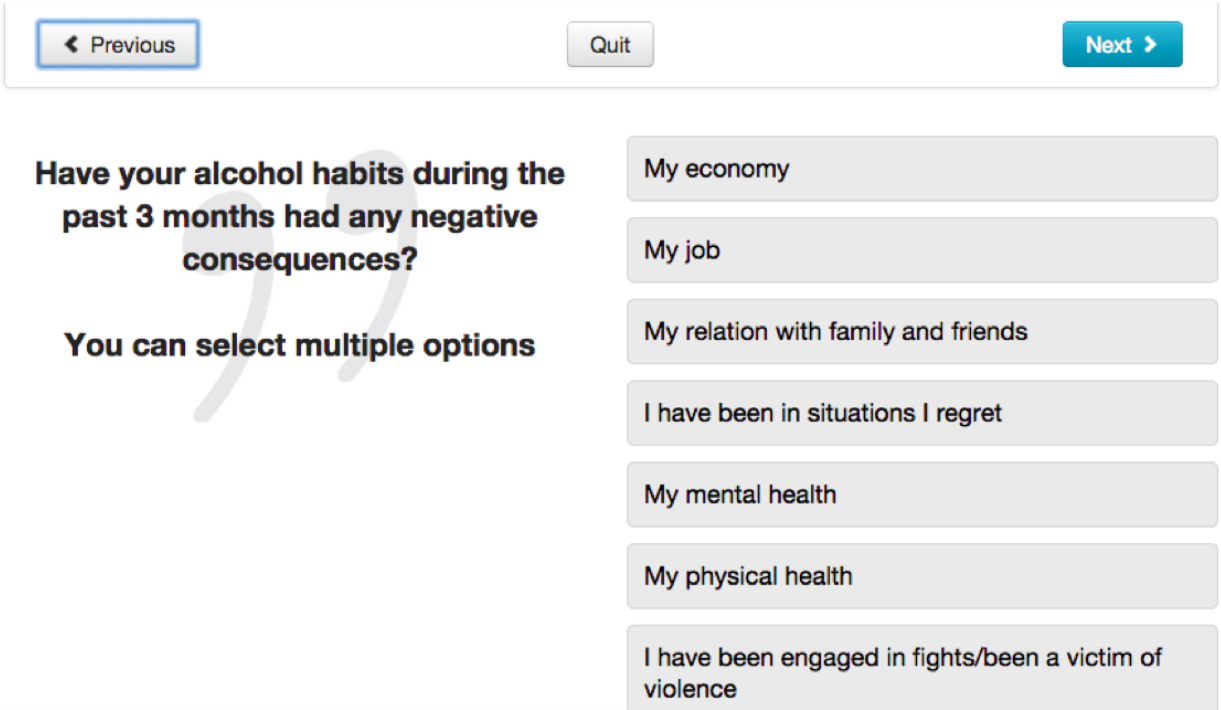


Figure 3. Screenshot of AMADEUS: assessment of negative consequences.



XSL•FO
RenderX

Sample Size

The marginal costs involved in delivering online interventions to large numbers of participants in both routine practice and in scientific studies are low and much lower than other brief interventions after the developmental costs are met [21-22]. Therefore, even very small effects are likely also to be highly cost effective above the basic threshold cost involved in providing the service. These observations led us to believe that the sample should be as large as possible in order to detect very small effects. To assist study planning, we undertook an illustrative power calculation. To detect an effect size of d=0.1 standard deviations between the 2 groups with 5% significance level and 80% power required 1600 individuals analyzed per group. Assuming a follow-up rate of 50%, we aimed at recruiting 3200 individuals per group (ie 6400 in total). We had no data on the number of screen positives who might be willing to participate in the trial, but assumed approximately 70% would do so, meaning that we would need to identify approximately 8000 hazardous and harmful drinkers. In order to identify these number of participants, we needed to send emails to approximately 40,000 students with an average response rate of 40% (ie, n=16,000) and a 50% prevalence rate. We could not be confident of these estimates because, for example, patterns of email use vary considerably between colleges, being compulsory in some institutions and rarely used in others. Therefore, we decided to undertake the study in 9 colleges/universities with a total student enrollment of 54,507 students.

Outcome Evaluation

This study used a single 2-month follow-up assessment interval, after which the control group gained access to the intervention. Thus, this study provides direct information only on the short-term effects of the intervention, although we stated a priori that if we found no short-term effects, we would not expect any longer-term effects [18].

The primary outcome was total weekly alcohol consumption. This was computed as the sum of alcohol consumed in standard drinks for each of the 7 days in a typical week, with data on each day of the week provided separately. Secondary outcomes were the proportions of students still drinking above national guidelines [25], frequency of drinking (number of days per week), quantity of drinks per drinking day, frequency of HED as defined in the screening question, highest estimated blood alcohol concentration (eBAC), and motivation to change.

Because there was no research assessment at study entry, information at this point was restricted to university, term, time from sending of invitation email to consent, and the frequency of HED from the screening question. At follow-up, we obtained additional information that was not possible, or not likely, to have been altered or altered differentially during the study period and which we used to examine equivalence: age, gender, weight, faculty of study, domestic status, and language used to answer the assessment and feedback language (Swedish or English). We also used measures of engagement with the study (device used, number of follow-up emails sent, and elapsed time before follow-up was completed).

Statistical Methods

All outcomes were compared between randomized groups under the intention-to-treat principle (ie, including all randomized individuals in their originally randomized groups). The characteristics of responders at follow-up were compared between groups using chi-square tests or Fisher's exact test for comparison of proportions and Student's t test for comparison of means. Wilcoxon rank sum test was used to compare groups regarding time to follow-up. A linear trend test was applied to detect a possible trend in proportion of responses in relation to the number of email reminders before responding. Continuous outcome measures were assessed for skewness by visual inspection of histograms and Q-Q plots. Skewed continuous outcome were analyzed by negative binomial regression, and results are reported as percent reduction. Drinking above national guidelines was analyzed by logistic regression and reported as percentage reduction in odds. Frequency of HED occasions was analyzed by ordered logistic regression and reported as a percentage reduction in odds for exceeding any level. All regression analyses were performed first unadjusted and then adjusted for frequency of HED at baseline, age, university, and gender using the first 2 as continuous variables (thus allowing for dependence between individuals in the same university); the adjusted analysis was specified a priori as the primary result. A sensitivity analysis excluded 3 outliers in the follow-up assessment with extreme reported weekly consumption values.

Missing outcome data were initially handled by a complete-case analysis assuming that the data were missing at random (MAR). If data were not MAR, then nonresponders differed systematically from responders and early responders were likely to differ systematically from late responders, who were likely to be more similar to nonresponders [23]. Therefore, we explored the plausibility of the MAR assumption by regressing the primary outcome on the number of follow-up emails needed before an individual responded using a negative binomial regression in responders: a significant association would cast doubt on the MAR assumption. To allow for the possibility of data being missing not at random, we fitted the repeated attempts model of Jackson et al [27]. This model was not available in standard software for negative binomial regression, so we applied it to a linear regression of log (alcohol consumption + k), where k=24 units/week was chosen to eliminate skewness.

Tests for whether the intervention effect was modified by frequency of HED at baseline, age, university, and gender were undertaken for the primary outcome only and the first 2 were used as continuous variables. A post hoc sensitivity analysis accounted for possible heterogeneity between universities of treatment effects on weekly alcohol consumption using a 2-stage approach. The treatment effects on weekly alcohol consumption were first estimated in each university separately by negative binomial regression (adjusted for frequency of HED at baseline, age, and gender using the first 2 as continuous variables) and were then combined in a random effects meta-analysis.

XSL•FO
RenderX

Results

Study Population Characteristics

Figure 4 depicts the flow of participants from the invitation to the follow-up. In total, 1605 risky (both hazardous and harmful) drinkers agreed to participate in the study and were randomized to the intervention arm (n=825) or control arm (n=780). In Table 1, the intervention and control groups are compared at baseline for frequency of HED, term, and time to consent. Some of the smaller participating college/universities only managed to recruit a few risky drinkers. For one college (Gävle), none of the 4 recruited participants were randomized to the control group, so this college was necessarily excluded from the remaining analyses including college. Two large universities (Lund and Uppsala) contributed approximately two-thirds of all participants. There were no differences between the intervention and control groups with regard to baseline characteristics.

Table 1. Comparison of groups at baseline (N=1605).

Baseline data	Intervention (n=825)	Control (n=780)
HED [a] occasions, n (%)		
2-3 times a month	434 (52.6)	422 (54.1)
1-2 times a week	350 (42.4)	323 (41.4)
≥3 times a week	41 (5.0)	35 (4.5)
University, n (%)		
Blekinge	23 (2.8)	31 (4.0)
Linné	85 (10.3)	65 (8.3)
Malmö	78 (9.5)	57 (7.3)
Lund	293 (35.5)	275 (35.3)
Gävle	4 (0.5)	0 (0.0)
Halmstad	27 (3.3)	38 (4.9)
Mälardalen	21 (2.5)	27 (3.5)
Skövde	24 (2.9)	22 (2.8)
Uppsala	270 (32.7)	265 (34.0)
Term, n (%)		
2	352 (42.7)	306 (39.2)
4	263 (31.9)	251 (32.2)
6	210 (25.5)	223 (28.6)
Time to consent (hours), median (IQR)	93 (4-200)	102 (5-190)

[a] HED: Heavy episodic drinking assessed by the question “How often, during the past 3 months, have you consumed 4 (women) or 5 (men) standard drinks on 1 occasion?”

The control group were much more likely to participate at follow-up (67.8% vs 49.0%, *P*<.001). Table 2 compares the characteristics of these responders in the intervention and control group at follow-up. None of the characteristics unlikely to have been altered since baseline differed between the 2 groups. There was a statistically significant decrease in the proportion of responses in relation to the number of email reminders before responding (*P*=.02).

XSL•FO
RenderX

Table 2. Comparison of groups at follow-up[a] (N=931).

Characteristics	Intervention (n=402)	Control (n=529)	*P* [b]
Characteristics unlikely to have changed since baseline			
Gender, n (%) [c]			
Male	198 (49.3)	276 (52.2)	.38
Female	204 (50.7)	253 (47.8)	
Age (years), n (%) [c]			
<18	62 (15.4)	81 (15.3)	.60
18-20	271 (67.4)	366 (69.2)	
21-25	49 (12.2)	65 (12.3)	
26-30	20 (5.0)	17 (3.2)	
Faculty of study, n (%) [c]			
Science and engineering	128 (31.8)	168 (31.8)	.90
Humanities	213 (53.0)	286 (54.1)	
Medical	61 (15.2)	75 (14.2)	
Language used, n (%) [c]			
Swedish	387 (96.3)	511 (96.6)	.79
English	15 (3.7)	18 (3.4)	
Domestic status, n (%)			
Living alone without kids at home	267 (66.4)	360 (68.1)	.06
Living alone with kids at home	3 (0.7)	1 (0.2)	
Living with somebody without kids	74 (18.4)	101 (19.1)	
Living with somebody with kids	15 (3.7)	6 (1.1)	
Have a partner but not living together	43 (10.7)	61 (11.5)	
Domestic status (3 categories), n (%)			
Living alone	270 (67.2)	361 (68.2)	.75
Living with somebody	89 (22.1)	107 (20.2)	
Have a partner but not living together	43 (10.7)	61 (11.5)	
Weight (kg), mean (SD)[d]	71.53 (13.35)	70.92 (12.24)	
Characteristics specific to follow-up			
Device used, n (%)			
Mobile phone	85 (21.1)	129 (24.4)	.39
Laptop	302 (75.1)	376 (71.1)	
Tablet	15 (3.7)	24 (4.5)	
Time to follow-up (hours), median (IQR)	71 (279)	48 (207)	0.31
Number of follow-up emails before response, n (%)			
1	222 (55.2)	315 (59.6)	.07[e]
2	82 (20.4)	103 (19.5)	
3	51 (12.7)	70 (13.2)	
4	27 (6.7)	23 (4.3)	
5	20 (5.0)	18 (3.4)	

[a] Without university of Gävle; this is the population used for the primary analyses in Table 3.

XSL•FO
RenderX

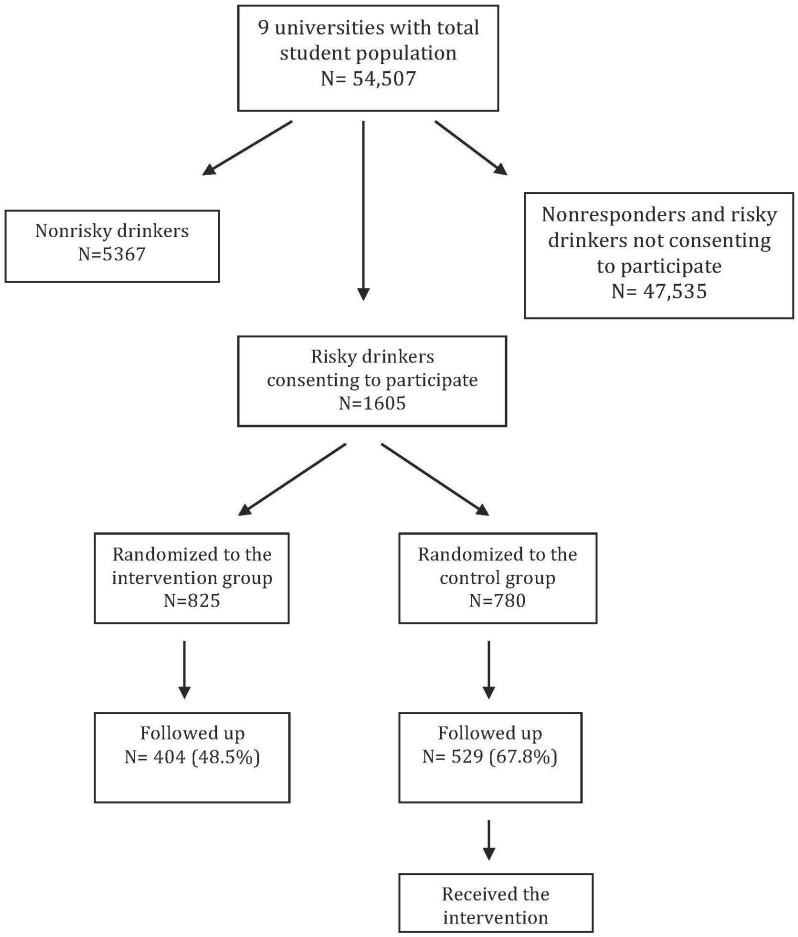
[b] All *P*-values were provided for heterogeneity except for variable number of follow-up emails. Determined using chi-square test (gender, age, faculty, language, domestic status 3 categories), Fisher exact test (domestic status), Student's *t* test (weight), Wilcoxon rank sum test (time to follow-up), or linear trend test (number of follow-up emails).

[c] Regarded as baseline variables in the analysis.

[d] Intervention: n=401; control: n=528.

[e] Trend test.

Figure 4. Flowchart of the AMADEUS study.



Main Findings

The main outcomes are displayed in Table 3. Small differences indicative of lower risk consumption were observed for the intervention group for all primary and secondary outcomes. However, none of these differences were statistically significant in the primary adjusted analyses.

XSL•FO
RenderX

Table 3. Trial outcomes (n=931).

Outcomes	Intervention (n=402)	Control (n=529)	% Reduction in mean or odds[a]			
			Unadjusted		Adjusted[b]	
			% (95% CI)	*P*	% (95% CI)	*P*
Primary outcome						
Weekly alcohol consumption (g/week), mean (SD)	113.4 (81.1)	120.8 (86.4)	6% (–3%, 14%)	.18	6% (–2%, 13%)	.13
Secondary outcomes						
Proportion drinking above national guidelines,[c] n (%)	370 (92.0)	492 (93.0)	13% (–42%, 47%)	.58	5% (–59%, 43%)	.85
Frequency of drinking (days/week), mean (SD)	2.30 (1.52)	2.34 (1.53)	2% (–7%, 10%)	.67	1% (–8%, 9%)	.85
Number of drinks per drinking day,[d] mean (SD)	4.5 (2.6)	4.7 (2.7)	5% (–2%, 12%)	.14	4% (–3%, 11%)*P*	.23
Frequency of HED occasions, n (%)			16% (–7%, 33%)	.17	14% (–10%, 33%)	.24
Never	11 (2.7)	8 (1.5)				
<1 time a month	22 (5.5)	29 (5.5)				
Approximately once a month	64 (15.9)	82 (15.5)				
2-3 times a month	169 (42.0)	207 (39.1)				
1-2 times a week	128 (31.8)	187 (35.3)				
≥3 times a week	8 (2.0)	16 (3.0)				
Highest eBAC, mean (SD)	1.16 (1.08)	1.31 (1.14)	11% (0%, 21%)	.05	11% (–1%, 20%)	.06
Motivation to change, [e] n (%)			4% (–21%, 25%)	.71	2% (–24%, 23%)	.86
I have had no thoughts about decreasing	175 (43.6)	225 (42.6)				
I have thought about decreasing, but I am not thinking about it right now	87 (21.7)	108 (20.5)				
I am thinking about how I will decrease	28 (7.0)	50 (9.5)				
I have started decreasing	105 (26.2)	135 (25.6)				
I have tried to decrease, but failed	6 (1.5)	10 (1.9)				

[a] Of intervention compared to control. Reduction in mean by negative binomial regression: weekly alcohol consumption, frequency of drinking, number of drinks per drinking day, and highest eBAC; reduction in odds by logistic regression: proportion drinking about national guidelines; reduction in odds of exceeding any cutoff by ordered logistic regression: frequency of HED occasions, motivation to change.

[b] Adjusted for frequency of heavy episodic drinking at baseline, age, university, and gender, using the first 2 as continuous variables.

[c] Risky drinker: heavy episodic drinking (HED) >once per month and/or total weekly consumption >14 standard drinks (men) or 9 (women).

[d] Intervention: n=395; control: n=523.

[e] Intervention: n=401; control: n=528.

Additional Analyses

Effect modification analyses did not reveal any statistically significant findings. However, the interaction between university and randomized group had a *P*-value of .07. Therefore, we explored this in a post hoc sensitivity analysis allowing the treatment effect to vary by university. The random effects meta-analysis (Figure 5) showed a 7% reduction in weekly alcohol consumption that was not statistically significant (95% CI –16% to 4%, *P*=.20). The confidence intervals of the analysis adjusted by cluster were somewhat wider than in the unadjusted analysis.

We also considered the statistically significant between-group difference for Uppsala (n=365), where weekly alcohol consumption was approximately 14% lower (95% CI –23% to –4%, *P*=.009 adjusted) in the intervention group than the control

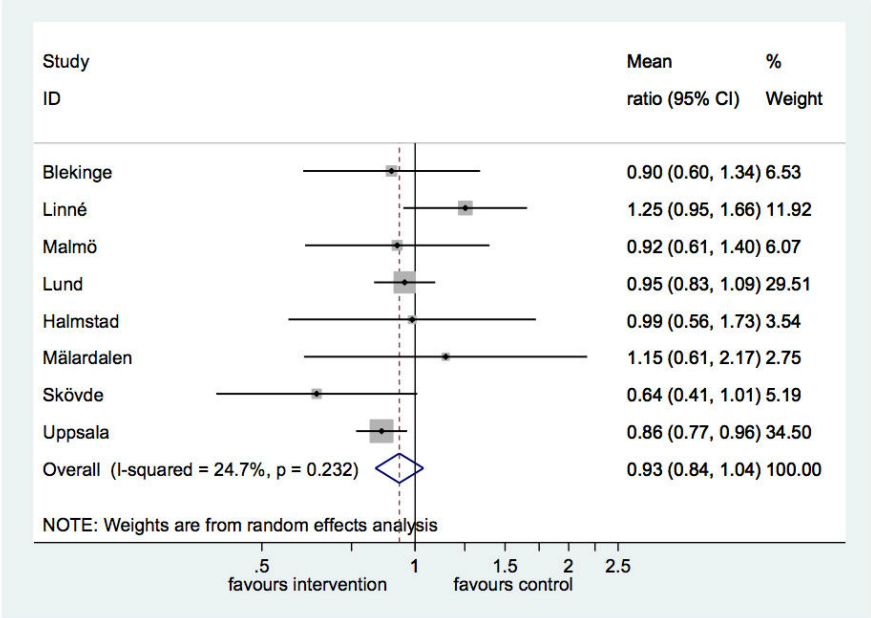
XSL•FO
RenderX

group. Further analysis of 3 secondary outcomes for Uppsala University also showed a significant difference in number of drinks per drinking day with 13% fewer (95% CI –23% to 3%, *P*=.01 adjusted), but no significant differences for frequency of drinking and highest eBAC.

In the assessment for skewness of the continuous variables, we found 1 outlier in the treatment group (with weekly alcohol consumption of 1044 g/week) and 2 outliers in the control group (with weekly alcohol consumption of 1128 and 1524 g/week). The maximum reported weekly alcohol consumption of those not excluded was 552 g/week in the treatment group and 456 g/week in the control group. Therefore, we performed a sensitivity analysis without these outliers. In this analysis, which was not specified a priori, the between-group difference in the primary outcome, weekly alcohol consumption, in the primary adjusted analysis, crossed the conventional threshold for statistical significance (8% reduction, 95% CI –15% to 0%, *P*=.049 adjusted; 10% reduction, 95% CI –17% to –1%, *P*=.02 unadjusted). No statistically significant differences were seen for the secondary outcomes in the primary adjusted analyses, although eBAC was statistically significant in the unadjusted analysis (11% reduction, 95% CI 0% to 21%, *P*=.047) but did not meet significance in the adjusted analysis (11% reduction, 95% CI –21% to 1%, *P*=.06).

The preceding analyses assumed the data were MAR. There was no statistically significant association between the primary outcome and the number of email reminders before answering the follow-up (*P*=.71), so the data are consistent with the MAR assumption. Another post hoc analysis assessed time to consent and found no association and thus no evidence that data were not MAR. Analyses using the repeated attempts model and linear regression suggested an intervention effect of a 10% reduction (95% CI –30% to 10%).

Figure 5. Forest plot of ratio of means in weekly alcohol consumption comparing intervention to control.



Discussion

The study found no strong evidence of short-term effectiveness of the Swedish national system of proactive online alcohol intervention for university and college students. However, inspection of the confidence intervals for the primary outcome (in Table 3) reveals that this study does not rule out an intervention effect of up to 13% reduction in total weekly alcohol consumption. The sensitivity analysis excluding outliers suggests an intervention effect on reduced total weekly alcohol consumption, although the statistical significance attained in that analysis should be treated with caution because that particular analysis was not prespecified. We have no reason to anticipate later occurring effects because brief intervention effects generally wane with time [5] and the short-term nature of this evaluation study is important to note in interpreting study findings. The best estimate of the intervention effect in Uppsala is larger (14% reduction in total weekly alcohol consumption compared to 6% across Sweden as a whole) and is statistically significant. This finding should be interpreted as hypothesis generating because, although prespecified, the differences in intervention effect across the universities as a whole were not statistically significant.

We did not have statistical power to detect the effect size that we believed was worth obtaining in the planning of this study and this is a clear study limitation. We succeeded in recruiting only one-quarter of our target sample size and the best estimate of the effect obtained on the primary outcome (a 6% reduction in alcohol consumption) is of clear public health significance. For example, it is very close to the size of effects considered appropriate for the implementation of face-to-face brief intervention programs [28] and measures to increase the price of alcohol [29]. Although we set out to evaluate a national system, we only managed to recruit 3% of all individuals invited, although many of those who did not respond will have simply

XSL•FO
RenderX

ignored the email invitation as not relevant to them (eg, because they did not drink much or at all).

The contrast with AMADEUS-1 in rates of participation is very striking. Conventional trial recruitment in AMADEUS-2 resulted in less than 1000 hazardous and harmful drinkers providing follow-up data from more than 54,000 initial invitations to consider participation in 9 universities and colleges. In AMADEUS-1, more than 1500 hazardous and harmful drinkers in 2 of the 3 arms provided follow-up data after completing baseline assessments, among approximately 7800 providing follow-up data from less than 15,000 targeted for study in 2 universities. Even allowing for differences in use of email between universities, the conventional study design with informed consent clearly impacts on participation in detrimental ways. However, other differences between the studies should be borne in mind. AMADEUS-2 provides an intention-to-treat evaluation of effects among hazardous and harmful drinkers, whereas a per-protocol analysis only was possible in AMADEUS-1 for this group due to the nature of the study design. Participation rates in AMADEUS-2 nevertheless expose the limitations of unblinded conventional trials designed to detect small effects of public health significance; it is not possible to undertake a fully powered study among student risky drinkers in Sweden, a country of approximately 10 million people. The external validity of the current findings, bearing in mind the low participation rate, warrants careful consideration.

Reliably detecting small effects is challenging and subject to the play of chance, and likely also to be influenced by a number of contextual factors that are difficult to capture. For example, the timing of follow-up within the academic year might be relevant because campus activities involving alcohol may both vary and influence study findings. Initial follow-up in this study was undertaken as exams approached.

The extent of differential attrition provides additional reasons for avoiding strong conclusions because the potential for selection bias exists, even though our analysis revealed no evidence to contradict our MAR assumption in relation to the missing data. It is also important to note that this intervention is not designed to meet the needs of problem drinkers [30] and online interventions extending over several sessions or contacts and/or person-to-person interactions are likely to be needed.

Previous online alcohol studies among college and university students have shown mixed results [6,7,29]. Apart from AMADEUS-1, there have been few randomized studies capable of dismantling components of effective interventions in this or similar populations. The content of this Swedish intervention is broadly similar to the THRIVE intervention evaluated in one university in Australia [31] and the New Zealand e-SBINZ trials intervention in comprising normative feedback, criterion feedback (on guidelines), and brief advice [32]. The effects are also broadly similar, for example with THRIVE showing a 17% reduction in alcohol consumption in comparison with controls after 1 month and decreasing to 11% after 6 months [31]. Some effects among Maoris were somewhat larger, although effects among non-Maori in the parallel e-SBINZ trial were smaller [32,33]. AMADEUS-1 [16] found that feedback added little to the effects of assessment and detailed investigations of intervention content are urgently needed in order to ascertain whether it may be possible to develop novel interventions or intervention components capable of larger effects than have been identified to date. This underdevelopment of online study is similar to that which pertains for face-to-face brief interventions [34].

The study has attempted to evaluate the effectiveness of an existing national service provision across a sizable number of institutions. Although it does not provide strong evidence of benefit, it is unclear how far the existing intervention should be redesigned, if at all, on the basis of these findings and those of AMADEUS-1. Further research and development work is needed, particularly in light of the low costs involved and the consequent likelihood of high levels of cost-effectiveness being associated with even very small effects. The key research challenge is to robustly identify and control for biases that interfere with reliable estimation of small effects [24,25,27-35]. Randomized controlled trials that directly compare the performance of the existing online intervention with novel candidates to augment or replace it are appropriate in this situation. Further investment in such work should not detract from the need for other population-level alcohol interventions, such as increasing price, better controlling availability, and restricting marketing to change the cultural acceptability of heavy drinking. The existing evidence suggests that these types of alcohol policies are most likely to be effective and we do not know whether or how far individual-level interventions in whole populations such as that evaluated here may enhance the anticipated effects [36].

Acknowledgments

The study was funded by the Swedish Council for Working Life and Social Research (FAS, in Swedish; Grant number 2010-0024) and by a Wellcome Trust Research Career Development fellowship in Basic Biomedical Science (WT086516MA) to JM. IW was supported by the Medical Research Council (Unit Program number: U105260558).

Authors' Contributions

JM and PB had the original idea for the study, obtained funding, and led its design. PB had overall responsibility for study implementation. MB did all computer programming associated both with intervention delivery and study data collection. IW and NK led on statistical aspects of this study, with inputs from JM. PB wrote the first draft of the manuscript to which all authors contributed revisions. All authors read and approved the final manuscript.

XSL•FO
RenderX

Conflicts of Interest

PB and MB own the company that developed the online intervention used in this study and that also develops and distributes computerized lifestyle interventions. None of the other authors have any conflicts to declare.

Multimedia Appendix 1

CONSORT-EHEALTH Checklist V 1.6.2 [37].

[PDF File (Adobe PDF File), 139KB - jmir_v17i7e170_app1.pdf]

References

1. Andersson A, Wiréhn A, Olvander C, Ekman DS, Bendtsen P. Alcohol use among university students in Sweden measured by an electronic screening instrument. BMC Public Health 2009;9:229 [FREE Full text] [doi: 10.1186/1471-2458-9-229] [Medline: 19594906]
2. Bendtsen P, Johansson K, Akerlind I. Feasibility of an email-based electronic screening and brief intervention (e-SBI) to college students in Sweden. Addict Behav 2006 May;31(5):777-787. [doi: 10.1016/j.addbeh.2005.06.002] [Medline: 15996827]
3. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, Fowkes Francis Gerry R, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec 15;380(9859):2224-2260 [FREE Full text] [doi: 10.1016/S0140-6736(12)61766-8] [Medline: 23245609]
4. McCambridge J, McAlaney J, Rowe R. Adult consequences of late adolescent alcohol consumption: a systematic review of cohort studies. PLoS Med 2011;8(2):e1000413 [FREE Full text] [doi: 10.1371/journal.pmed.1000413] [Medline: 21346802]
5. Moyer A, Finney JW, Swearingen CE, Vergun P. Brief interventions for alcohol problems: a meta-analytic review of controlled investigations in treatment-seeking and non-treatment-seeking populations. Addiction 2002 Mar;97(3):279-292. [Medline: 11964101]
6. Carey KB, Scott-Sheldon LA, Elliott JC, Garey L, Carey MP. Face-to-face versus computer-delivered alcohol interventions for college drinkers: a meta-analytic review, 1998 to 2010. Clin Psychol Rev 2012 Dec;32(8):690-703 [FREE Full text] [doi: 10.1016/j.cpr.2012.08.001] [Medline: 23022767]
7. Donoghue K, Patton R, Phillips T, Deluca P, Drummond C. The effectiveness of electronic screening and brief intervention for reducing levels of alcohol consumption: a systematic review and meta-analysis. J Med Internet Res 2014;16(6):e142 [FREE Full text] [doi: 10.2196/jmir.3193] [Medline: 24892426]
8. Khadjesari Z, Murray E, Hewitt C, Hartley S, Godfrey C. Can stand-alone computer-based interventions reduce alcohol consumption? A systematic review. Addiction 2011 Feb;106(2):267-282. [doi: 10.1111/j.1360-0443.2010.03214.x] [Medline: 21083832]
9. Murray E, Khadjesari Z, White IR, Kalaitzaki E, Godfrey C, McCambridge J, et al. Methodological challenges in online trials. J Med Internet Res 2009;11(2):e9 [FREE Full text] [doi: 10.2196/jmir.1052] [Medline: 19403465]
10. Elliott JC, Carey KB, Bolles JR. Computer-based interventions for college drinking: a qualitative review. Addict Behav 2008 Aug;33(8):994-1005 [FREE Full text] [doi: 10.1016/j.addbeh.2008.03.006] [Medline: 18538484]
11. Carey KB, Scott-Sheldon LA, Elliott JC, Bolles JR, Carey MP. Computer-delivered interventions to reduce college student drinking: a meta-analysis. Addiction 2009 Nov;104(11):1807-1819 [FREE Full text] [doi: 10.1111/j.1360-0443.2009.02691.x] [Medline: 19744139]
12. Bendtsen P, McCambridge J, Bendtsen M, Karlsson N, Nilsen P. Effectiveness of a proactive mail-based alcohol Internet intervention for university students: dismantling the assessment and feedback components in a randomized controlled trial. J Med Internet Res 2012;14(5):e142 [FREE Full text] [doi: 10.2196/jmir.2062] [Medline: 23113955]
13. McCambridge J, Kypri K. Can simply answering research questions change behaviour? Systematic review and meta analyses of brief alcohol intervention trials. PLoS One 2011;6(10):e23748 [FREE Full text] [doi: 10.1371/journal.pone.0023748] [Medline: 21998626]
14. Cunningham JA, Kypri K, McCambridge J. The use of emerging technologies in alcohol treatment. Alcohol Res Health 2011;33(4):320-326 [FREE Full text] [Medline: 23580017]
15. McCambridge J, Bendtsen P, Bendtsen M, Nilsen P. Alcohol email assessment and feedback study dismantling effectiveness for university students (AMADEUS-1): study protocol for a randomized controlled trial. Trials 2012;13:49 [FREE Full text] [doi: 10.1186/1745-6215-13-49] [Medline: 22540638]
16. McCambridge J, Bendtsen M, Karlsson N, White IR, Nilsen P, Bendtsen P. Alcohol assessment and feedback by email for university students: main findings from a randomised controlled trial. Br J Psychiatry 2013 Nov;203(5):334-340 [FREE Full text] [doi: 10.1192/bjp.bp.113.128660] [Medline: 24072758]

XSL•FO
RenderX

17. McCambridge J, Kypri K, Bendtsen P, Porter J. The use of deception in public health behavioral intervention trials: a case study of three online alcohol trials. Am J Bioeth 2013;13(11):39-47 [FREE Full text] [doi: 10.1080/15265161.2013.839751] [Medline: 24161181]
18. McCambridge J, Bendtsen M, Karlsson N, White IR, Bendtsen P. Alcohol assessment & feedback by e-mail for university student hazardous and harmful drinkers: study protocol for the AMADEUS-2 randomised controlled trial. BMC Public Health 2013;13:949 [FREE Full text] [doi: 10.1186/1471-2458-13-949] [Medline: 24456668]
19. Babor TF, Higgins-Biddle JC, Saunders JB, Monteiro MG. AUDIT-Alcohol Use Disorders Identification Test: Guidelines for Use in Primary Care. Geneva: World Health Organization; 2001.
20. Canagasaby A, Vinson DC. Screening for hazardous or harmful drinking using one or two quantity-frequency questions. Alcohol Alcohol 2005;40(3):208-213 [FREE Full text] [doi: 10.1093/alcalc/agh156] [Medline: 15797883]
21. Johnson JA, Lee A, Vinson D, Seale JP. Use of AUDIT-based measures to identify unhealthy alcohol use and alcohol dependence in primary care: a validation study. Alcohol Clin Exp Res 2013 Jan;37 Suppl 1:E253-E259. [doi: 10.1111/j.1530-0277.2012.01898.x] [Medline: 22834916]
22. Wicki M, Kuntsche E, Gmel G. Drinking at European universities? A review of students' alcohol use. Addict Behav 2010 Nov;35(11):913-924. [doi: 10.1016/j.addbeh.2010.06.015] [Medline: 20624671]
23. Walters ST, Vader AM, Harris TR, Jouriles EN. Reactivity to alcohol assessment measures: an experimental test. Addiction 2009 Aug;104(8):1305-1310 [FREE Full text] [doi: 10.1111/j.1360-0443.2009.02632.x] [Medline: 19624323]
24. McCambridge J, Kypri K, Elbourne D. In randomization we trust? There are overlooked problems in experimenting with people in behavioral intervention trials. J Clin Epidemiol 2014 Mar;67(3):247-253 [FREE Full text] [doi: 10.1016/j.jclinepi.2013.09.004] [Medline: 24314401]
25. Andréasson S, Allebeck P. Alcohol and Health [Alkohol och Hälsa]. Stockholm: The Swedish National Institute of Public Health; 2005:91-72573.
26. Linköping University. Questionnaire URL: https://livsstilsanalys.alexit.se/?managerId=106751&questionSet=alcohol_student [WebCite Cache ID 6ZfgySC1P]
27. Jackson D, White IR, Leese M. How much can we learn about missing data?: an exploration of a clinical trial in psychiatry. J R Stat Soc Ser A Stat Soc 2010 Jul;173(3):593-612 [FREE Full text] [doi: 10.1111/j.1467-985X.2009.00627.x] [Medline: 20711246]
28. Purshouse RC, Brennan A, Rafia R, Latimer NR, Archer RJ, Angus CR, et al. Modelling the cost-effectiveness of alcohol screening and brief interventions in primary care in England. Alcohol Alcohol 2013;48(2):180-188 [FREE Full text] [doi: 10.1093/alcalc/ags103] [Medline: 23015608]
29. Purshouse RC, Meier PS, Brennan A, Taylor KB, Rafia R. Estimated effect of alcohol pricing policies on health and health economic outcomes in England: an epidemiological model. Lancet 2010 Apr 17;375(9723):1355-1364 [FREE Full text] [doi: 10.1016/S0140-6736(10)60058-X] [Medline: 20338629]
30. McCambridge J, Rollnick S. Should brief interventions in primary care address alcohol problems more strongly? Addiction 2014 Jul;109(7):1054-1058 [FREE Full text] [doi: 10.1111/add.12388] [Medline: 24433291]
31. Kypri K, Hallett J, Howat P, McManus A, Maycock B, Bowe S, et al. Randomized controlled trial of proactive web-based alcohol screening and brief intervention for university students. Arch Intern Med 2009 Sep 14;169(16):1508-1514. [doi: 10.1001/archinternmed.2009.249] [Medline: 19752409]
32. Kypri K, McCambridge J, Vater T, Bowe SJ, Saunders JB, Cunningham JA, et al. Web-based alcohol intervention for Māori university students: double-blind, multi-site randomized controlled trial. Addiction 2013 Feb;108(2):331-338 [FREE Full text] [doi: 10.1111/j.1360-0443.2012.04067.x] [Medline: 22925046]
33. Kypri K, McCambridge J, Cunningham JA, Vater T, Bowe S, De Graaf B, et al. Web-based alcohol screening and brief intervention for Māori and non-Māori: the New Zealand e-SBINZ trials. BMC Public Health 2010;10:781 [FREE Full text] [doi: 10.1186/1471-2458-10-781] [Medline: 21176233]
34. McCambridge J. Brief intervention content matters. Drug Alcohol Rev 2013 Jul;32(4):339-341 [FREE Full text] [doi: 10.1111/dar.12044] [Medline: 23819570]
35. Heather N. Interpreting null findings from trials of alcohol brief interventions. Front Psychiatry 2014;5:85 [FREE Full text] [doi: 10.3389/fpsyt.2014.00085] [Medline: 25076917]
36. Heather N. Can screening and brief intervention lead to population-level reductions in alcohol-related harm? Addict Sci Clin Pract 2012;7:15 [FREE Full text] [doi: 10.1186/1940-0640-7-15] [Medline: 23186309]
37. Eysenbach G, Consort- E. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

AUDIT: Alcohol Use Disorders Identification Test
eBAC: estimated blood alcohol concentration
HED: heavy episodic drinking

XSL•FO
RenderX

MAR: missing at random
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 11.11.14; peer-reviewed by N Johnson, M Boynton; comments to author 17.02.15; revised version received 25.03.15; accepted 28.04.15; published 09.07.15.

Please cite as:
Bendtsen P, Bendtsen M, Karlsson N, White IR, McCambridge J
Online Alcohol Assessment and Feedback for Hazardous and Harmful Drinkers: Findings From the AMADEUS-2 Randomized Controlled Trial of Routine Practice in Swedish Universities
J Med Internet Res 2015;17(7):e170
URL: http://www.jmir.org/2015/7/e170/
doi:10.2196/jmir.4020
PMID:26159179

©Preben Bendtsen, Marcus Bendtsen, Nadine Karlsson, Ian R White, Jim McCambridge. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 09.07.2015. 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

How to Increase Reach and Adherence of Web-Based Interventions: A Design Research Viewpoint

Geke DS Ludden[1*], PhD; Thomas JL van Rompay[1*], PhD; Saskia M Kelders[2*], PhD; Julia EWC van Gemert-Pijnen[2*], PhD

[1]Department of Design, Faculty of Engineering Technology, University of Twente, Enschede, Netherlands

[2]Center for eHealth Research and Disease Management, Department of Psychology, Health and Technology, University of Twente, Enschede, Netherlands

[*]all authors contributed equally

Corresponding Author:
Geke DS Ludden, PhD
Department of Design
Faculty of Engineering Technology
University of Twente
Horstring W254
de Horst
Enschede, 7522 LW
Netherlands
Phone: 31 534892920
Fax: 31 534893631
Email: g.d.s.ludden@utwente.nl

Abstract

Nowadays, technology is increasingly used to increase people's well-being. For example, many mobile and Web-based apps have been developed that can support people to become mentally fit or to manage their daily diet. However, analyses of current Web-based interventions show that many systems are only used by a specific group of users (eg, women, highly educated), and that even they often do not persist and drop out as the intervention unfolds. In this paper, we assess the impact of design features of Web-based interventions on reach and adherence and conclude that the power that design can have has not been used to its full potential. We propose looking at design research as a source of inspiration for new (to the field) design approaches. The paper goes on to specify and discuss three of these approaches: personalization, ambient information, and use of metaphors. Central to our viewpoint is the role of positive affect triggered by well-designed persuasive features to boost adherence and well-being. Finally, we discuss the future of persuasive eHealth interventions and suggest avenues for follow-up research.

(J Med Internet Res 2015;17(7):e172) doi:10.2196/jmir.4201

KEYWORDS

Web-based interventions; adherence; design for well-being; metaphors; personalization; ambient information

Introduction

Health Today

Our society faces severe problems when it comes to securing health for the public at large. First of all, unhealthy lifestyles, leading to diseases like obesity, are responsible for an unrelenting rise in health care costs. Secondly, long-term diseases like chronic obstructive pulmonary disease (COPD) and diabetes require life-long management of illness and attention from care professionals for a growing group of people. And thirdly, an aging population demands more intensive care. Next to (predominantly) physical health problems, a substantial number of people suffer from mental illnesses like burnout and stress disorders that likewise affect physical health in the long run.

Technology can play a role in stimulating people to take responsibility for their own health and well-being. Before we give examples of how technology can support people in taking this responsibility we will explain the relationship between health and well-being, and define the term well-being as we will use it throughout this paper.

Well-Being, Health, and Technology

Several researchers have found that there is a two-way relationship between well-being and health: health influences well-being and well-being itself influences health. Both physical

XSL•FO
RenderX

health and mental health influence well-being [1,2] and there are a number of correlations between well-being and physical health outcomes, such as improved immune system response, higher pain tolerance, increased longevity, cardiovascular health, slower disease progression, and reproductive health [3,4]. A meta-analysis even showed that emotional well-being predicts long-term prognosis of physical illness [5].

Different perspectives on the term well-being can be differentiated. Following Seligman's [6] well-being theory, we see well-being as a multi-componential concept comprising positive emotions, engagement, meaning, positive relationships, and accomplishment. We will further elaborate on how these different elements can contribute to well-being in the following sections.

As said, technology can play an important role in helping people to take responsibility for—and, for that matter, to give them more control over—their own health and well-being; and it can do so in several ways. Firstly, technology can play a role in helping people to manage their disease(s). Web-based platforms offer a way to communicate with health professionals and/or to manage health conditions. For example, people who suffer from diabetes can make use of apps that help them understand what makes blood sugar levels rise and fall, and thus how to control such fluctuations. Next to this, medical technology has been introduced that supports people to take physiological measurements at home and automatically send these to a medical professional. For example, people who need regular blood pressure monitoring can now use certified devices at home. For many people, this is a more acceptable way of continuous regular monitoring [7]. Furthermore, technology can provide health treatment; in the area of mental health, a variety of Web-based interventions have been shown to be effective in providing (guided) self-help therapy to reduce complaints, as in Barak et al [8], or to lead a more flourishing life. Chitarro and Vianello [9] explored how technology can support mindfulness exercises and found that technology-supported mindfulness may be more effective than traditional (paper-based) exercises. Technology can also play a role in helping people to lead a healthier lifestyle in order to prevent diseases. In this area, devices that track physical behavior and provide feedback are often combined with a range of mobile and Web-based interventions that help people to lose weight, to be more physically active, or to adopt a healthier diet.

Reach and Adherence: The Role of Design

The developments and examples above show that technology and, more specifically, Web-based interventions have a huge potential in (preventive) health and well-being. However, Web-based interventions often suffer from nonadherence: many people do not follow a treatment online as it was intended by the therapist. A systematic review of 83 Web-based interventions on lifestyle, chronic disease, and mental health found that, on average, around 50% of the participants adhere fully to an intervention [10]. This seems to reduce treatment effectiveness [11]. On top of this, most interventions reach a limited group only, while aiming for a broad audience. This does not have to be a problem; a strength of Web-based interventions may be that they can reach groups that are harder to reach by regular, face-to-face interventions, as seen in Postel et al [12] where e-therapy for problem drinking reached higher-educated females, while this group is underrepresented in regular care. However, several researchers have shown that current Web-based interventions often reach higher-educated women, as seen in several studies [13-17], while they fail to reach other groups in society. In some cases, these findings may have been influenced by the topic of the intervention under study, for example, the studies of Rothert et al [13], Binks and van Mierlo [14], and Kelders et al [15] were all aimed at weight management, a topic that may be of particular interest to higher-educated women. However, the overrepresentation of this specific group was also found for areas like alcohol abuse [12], depression [16], and smoking [17]; topics that seem at least relevant to many more groups in society. It seems there is something about current Web-based interventions themselves that attracts one specific group, and that in spite of their large potential, Web-based interventions miss out on helping the public at large. Kelders [10] argues that the design of interventions is an important factor for adherence and shows that there are persuasive features (eg, dialogue support, reminders and praise) when implemented in Web-based interventions that predict higher adherence. Moreover, in their review study on the effectiveness of eHealth interventions for physical activity and dietary behavior change, Norman et al [18] claim that interventions that feature interactive technologies in particular need to be refined to fully live up to their potential. In line with this claim, several researchers have proposed that the affective experience of persuasive technologies is the key to their effectiveness [19,20]. They reasoned that if the experience of, for example, a Web-based intervention is a pleasant one, people are inclined to keep using the intervention or to use the intervention again at a later point in time. In other words, a better design that is not only functionally effective but also desirable, compelling, and delightful could improve acceptance of, and adherence to, Web-based interventions. Introducing a "better" design to increase the affective experience of using a Web-based intervention could have a positive effect on well-being because it triggers positive emotions (the first element abbreviated as "P" in positive emotions, engagement, positive relationships, meaning, and accomplishment [PERMA]). However, this is a rather limited view on design's impact on well-being. In an effort to further specify how design can contribute to well-being, Pohlmeyer [21] constructed a design well-being matrix that specifies how design can have an impact on the different elements in Seligman's [6] well-being theory. As described, in this theory Seligman distinguishes five elements that can each contribute to a general feeling of well-being: positive emotions, engagement, positive relationships, meaning, and accomplishment. Pohlmeyer argues that design can take different roles that designers can use to intentionally design for well-being: source, symbol, enablement, and support. As examples, she describes how a product such as Paro the therapeutic robot seal, which is being used in care homes with dementia patients, can be the direct *source* of a *relationship* (R in PERMA) and how products can have an indirect effect on elements of PERMA by serving as a *symbol*. Think, for example, of a wedding ring (symbol of a relationship) or of a trophy (symbolizing achievement).

XSL•FO
RenderX

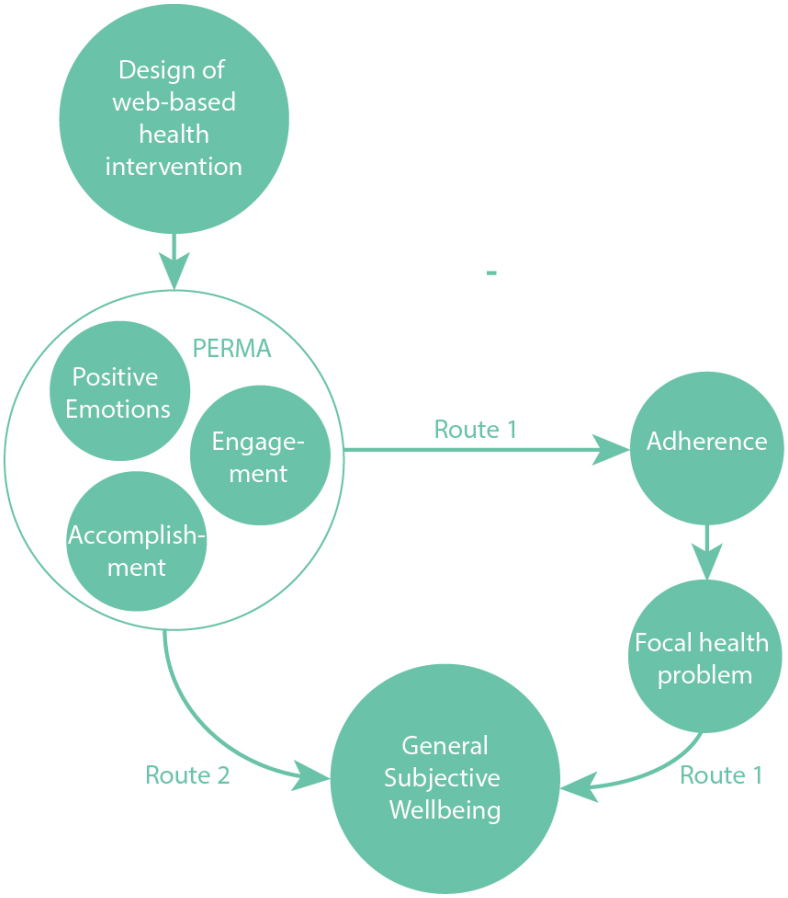
Influencing Well-Being via Two Routes

In the context of the design of a Web-based intervention, we would like to argue that design can have a positive effect on well-being following two different routes. First of all, a design aimed at a positive user experience by inducing, for example, positive emotion and/or engagement (P and E in PERMA) could positively influence well-being. Secondly, well-being during use could have an indirect effect on overall well-being because it can have a positive effect on adherence (ie, using the intervention as intended by the therapist). And better adherence to a Web-based intervention eventually has a positive effect on health, and thus on well-being. Next to this, successfully using the Web-based intervention could lead to a feeling of accomplishment (A in PERMA), again directly positively influencing well-being. Figure 1 shows how technology-supported health interventions can influence overall well-being via two routes.

The idea here is that, whereas the first route has a direct impact on the specific health problem (ie, people adhere and thus can manage and cure the focal health problem), the second route stimulates overall well-being by promoting states of mind that are conducive to a general feeling of well-being.

In this paper, our aim is not to test the relationships suggested in Figure 1. Rather, it should serve as an illustration that underlines the importance and the power of design. In the following sections, we will present a design research viewpoint to Web-based interventions and discuss new—to the field of designing Web-based interventions—design tools and approaches. We will explain and discuss how implementing these approaches could have a positive effect on reaching target groups and on adherence. To start, we will zoom in on Web-based health interventions and give a short overview of the types that are currently used. Next, we will illustrate with three examples how design could influence the experience of these interventions and contribute to an increased level of well-being when using such interventions. Finally, we discuss the future of Web-based health interventions and other technology-based, well-being interventions.

Figure 1. Schematic representation of how the design of a Web-based intervention can influence general subjective well-being following two different routes. Route 1 indicates the impact of design on adherence and thus on the focal health problem. Route 2 indicates how overall well-being is stimulated by elements of PERMA.



Web-Based Interventions

Overview

Web-based interventions can be categorized within the second generation of health interventions that make use of technology to transfer information [18]. Third generation interventions include mobile and remote devices such as mobile phones. We will include this third generation in some of our examples and in the discussion, but we will for now focus on the design of Web-based interventions because these form a widely used platform to implement technology-based health interventions.

Current Design of Web-Based Interventions

Design of Web-based interventions is often content driven and text based, and aimed at education, information, and goal setting

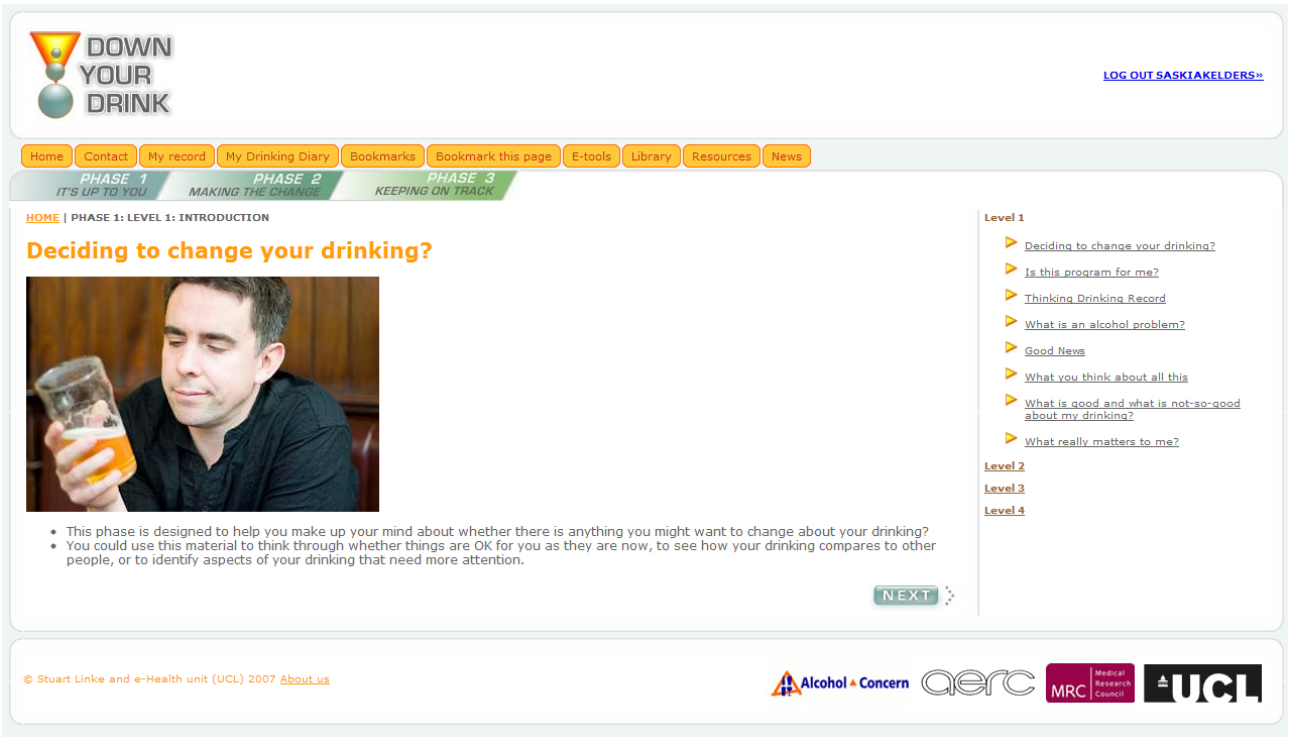
XSL•FO
RenderX

via modules or programs that have to be followed in a strict or fixed order. In many cases, content for a Web-based intervention was developed earlier as a fixed (offline) program that should later be “placed” on the website or mobile phone app. This often results in a text-driven, handout app, with the look and feel of a self-help book. Research on nonadherence indicates that reasons for dropout relate to dissatisfaction with the intervention, mismatch of goals of the intervention and those of the users, and low flexibility to adjust to different situations and user characteristics [15,22,23]. As a consequence, irritation rather than positive emotion (P in PERMA), and frustration rather than feelings of accomplishment (A in PERMA) could be triggered. Seen in this light, it should come as no surprise that low adherence is inherent to such interventions. However, it is precisely this "adherence" factor that is so essential for a positive effect of an intervention on health and well-being to transpire [11].

To further illustrate how low adherence may follow from ill-suited design, consider the Web-based intervention in Figure 2, aimed at reducing alcohol consumption. This intervention consists of six lessons, which are intended to be completed during 6 weeks [24,25]. Although some of the capabilities of technology are employed (eg, the provision of tailored feedback), the intervention is text based and technology seems to be mostly used as a medium to deliver the text. An analysis revealed that only 16.5% (of the first 10,000 registered users of the intervention) completed all 6 weeks of this intervention. A reason for nonadherence may well relate to a lack of attention to design and the resulting dissatisfying user experience. Limited reach of mobile and Web-based interventions could be explained by the fact that many of these interventions are aimed at a wide demographic group (eg, men and women, young and old, highly and less educated). Think, for example, of the range of track-and-trace systems aimed at keeping a healthy diet, losing weight, or increasing physical activity (eg, Lifesum, LosIt, MyFitnessPal). However, as we have argued in our introduction, most interventions reach a limited group only, mainly highly educated women—see, for example, Kelders et al [15]. In one way, this is a good thing, since this group in particular can be hard to reach with traditional treatment. However, this selective reach is not intended and, in many cases, seems to strengthen the "inverse care and information law" (ie, people in urgent need for care are the ones who are least likely to receive care [26,27]).

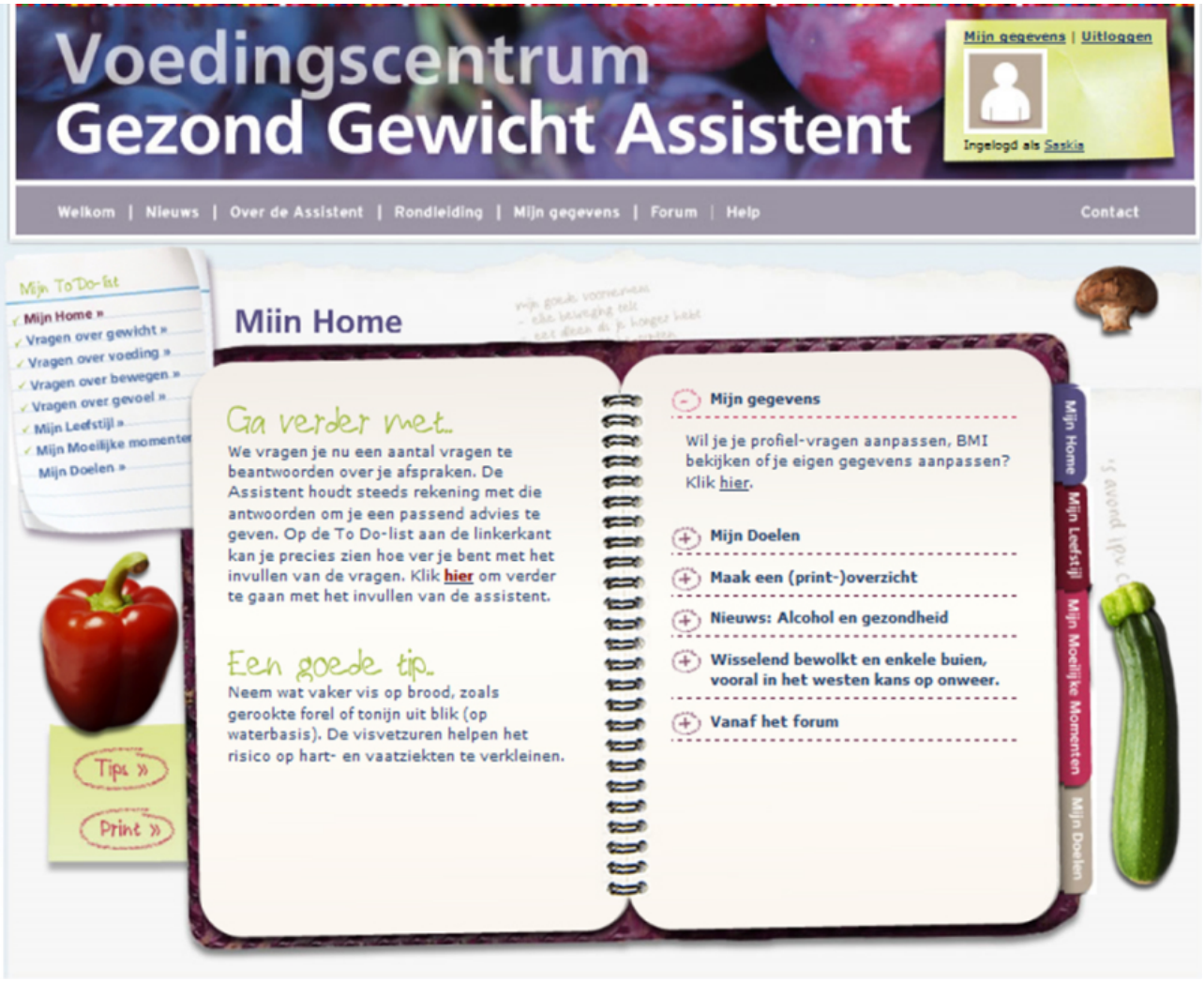
Consider another example of an intervention facing problems with reach and adherence (see Figure 3). The Healthy Weight Assistant is aimed at adults with healthy weight or who are slightly overweight. However, users were mainly female and highly educated [15]. Moreover, adherence to this intervention was as low as 3%, which the authors attribute to a mismatch between the goals of the intervention—long-term weight management—and the goals of the users—gain insight into their behaviors. This mismatch may have led to the low satisfaction and adherence found in the study.

Figure 2. The Web-based intervention, Down Your Drink.



XSL•FO
RenderX

Figure 3. The Web-based intervention, Healthy Weight Assistant.



Effect of Design on Adherence

There have been multiple attempts to investigate the influence of characteristics of interventions on adherence. Two qualitative systematic reviews investigated the influence of specific characteristics on adherence and indicated that increased personal relevance, an individualized approach (eg, involving tailored advice and feedback), and including clinicians are promising avenues [28]. In terms of PERMA, these characteristics are most directly related to engagement. Perceived relevance promotes engagement (E) and positive relationships (R), furthermore, including clinicians introduces a relationship (R). Additionally, peer support, counselor support, email/phone contact with visitors, and updates of the intervention website—direct means to trigger relationship (R) perceptions—resulted in improved reach [29]. Note that introducing such personal elements to trigger relationships is a costly and possibly time-consuming way to improve adherence. From a design point of view, it would be interesting to introduce relationship perceptions through design elements (eg, by color, layout, and typeface) connoting relationship-related meanings, such as involved, empathic, and dedicated. Alternatively, one could think of introducing relationships with digital avatars.

A third systematic review investigated the influence of persuasive technology features and characteristics of Web-based interventions [10], demonstrating the importance of increased interaction with a counselor, more frequent updates, and more extensive employment of dialogue support. This latter aspect in particular seems to provide a starting point for designing for well-being. The concept of dialogue support stems from the persuasive systems design model [30] and refers to supporting the interaction between user and system to facilitate progress toward goal fulfillment. These principles include social and cognitive prompts, such as praise, rewards, reminders, and suggestions.

There have also been studies that investigate the influence of one or more features on adherence in a single intervention. For example, telephone reminders have been shown to increase adherence to a Web-based treatment of social phobia without clinician guidance [31], and emailed messages have been shown to lead to a modest increase in usage of a disease prevention website by some adults [32]. Although these studies suggest a relationship between the design of an intervention and adherence, only a few of these interventions were actually, and purposefully, designed with adherence in mind [10]. This suggests that adapting Web-based interventions to promote adherence is done in an ad hoc manner or is considered as a task for the counselor involved in the intervention. Studies that *have* tried to design for adherence have mainly focused on

XSL•FO
RenderX

adding or adjusting features without explicitly considering the user experience as a means to improve reach and adherence.

Design Approaches to Increase Reach and Adherence

Overview

This short overview and discussion of Web-based interventions showed that while key works in the field are beginning to address reach and adherence challenges through increased interactivity and some possibilities for personalization, using a design strategy that effectively increases reach and/or adherence remains challenging. At the same time, technology offers many opportunities here; for example, it enables different means of communication simultaneously (eg, text, speech, video, and graphics) and provides access to situations and settings (eg, the bathroom) in which human persuaders would not be allowed in, or have no access to (eg, sensors in clothes) [33,19]. Therefore, we argue that for the design of Web-based interventions, it is worthwhile to look at how the field of design research is incorporating design strategies to design technology for well-being that delights people during use.

While seeking inspiration in the field of design research, it must be mentioned that a mere focus on expressive, creative design may result in technology that is likewise hard to understand and that has no fit with users' mental models. As such, creative designs can overshoot the mark because users do not experience such designs as supportive. On the other hand, the more traditional approach of carrying out a needs-and-demand assessment provides information about some functionalities of a design, but may not reveal that much about motivational cues for adherence and experience. Therefore, we propose to use other design approaches that go beyond both the mere "design creativity-based approaches" as well as requirements engineering.

In the following sections, we discuss three design strategies—personalization, use of ambient information, and use of metaphors—that have been particularly well explored in design research as strategies to attract and involve people. These can have a positive effect on how someone experiences a Web-based intervention, thereby increasing both well-being during use and overall well-being. Cooperation among experts from different disciplines and discussion along these lines may help in creating technology-based interventions that contribute to well-being by targeting the three dimensions of happiness as discussed by Seligman [34]: the pleasant life, the good life, and the meaningful life (or positive affect, flow, and meaning). Thus, technology should render Web-based apps more appealing, engaging, and fun (positive affect), stimulate flow by paving the way for smooth and intuitive interaction—for example, by providing ambient rather than in-your-face or, reversely, completely hidden information—and create meaning by giving users a sense of control, direction, and purpose.

Design Approach I: Personalization

In different application fields, researchers have shown that personalization of systems' functionalities and content can improve people's satisfaction with services and can increase users' efficiency and convenience. In PERMA terms, we would like to argue that personalization, foremost, promotes engagement (E) as it involves users and heightens personal relevance. Halko and Kientz [35] revealed significant relationships between personality and different types of persuasive technologies. Kaptein et al [36] measured susceptibility to persuasion and studied effects of tailored, persuasive text messages to reduce snacking. Results showed that tailored messages lead to a higher decrease in snacking consumption. In line with these findings, Lee et al [37] demonstrated that in the design of persuasive technologies for healthy eating, planning strategies worked differently depending on whether the participants had already adopted healthy dietary lifestyles.

These examples suggest that effects of design on adherence vary depending on users' personal needs and that personalization is therefore important, as it enables connecting to specific needs of different people. There are three ways to personalize a Web-based intervention and create larger reach. The first way is in line with the examples mentioned above and involves tailoring of messages or (persuasive) approach. Secondly, a designer can set out to design for a specific target group. To do this effectively, knowing what a specific group needs and wants, and what motivates them is essential. A designer can use this knowledge to inspire and direct the creative processes leading to an intervention that creates engagement for a specific group of people. An excellent example of a Web-based intervention that is aimed at a specific target group and was designed with the needs, habits, and desires of this target group in mind is the game Na-Aapje that was released by the Dutch Voedingscentrum (The Netherlands Nutrition Centre) a few years ago. Na-Aapje (loosely translated as little copy-cat) is a children's game that is designed to raise children's awareness of fruits and vegetables as healthy diet choices. The monkey in the game has to collect food items resulting in a higher overall score if many fruits and vegetables are collected (see Figure 4).

In this way, Na-Aapje links a, perhaps, not pleasurable (ie, eating fruits and vegetables) but necessary activity for its user—the child—to a more pleasurable one. Many children like to play computer games and by connecting to their preferred activities, adherence to this intervention (and awareness of the importance of eating fruits and vegetables) is probably increased.

A third way to personalize a Web-based intervention could be to design an intervention in such a way that it can be changed or set up at the start to match user preferences. This could be done at both the content level and at the system (design) level. From a psychological point of view, this user-controlled type of personalization is particularly interesting as it gives users a sense of control or dominance which may contribute to general well-being. For instance, in Mehrabian and Russell's [38] framework addressing emotional experiences in environmental settings, dominance is, next to pleasure and arousal, considered an essential factor in explaining approach-avoidance behavior, with higher degrees of control generally related to increases in approach behaviors—a notion similar to adherence in the online context—and hence, increased well-being.

XSL•FO
RenderX

Of further interest in this context is Averill's [39] discussion of "decisional control," defined as the degree to which a specific action results from choice among various alternatives. For instance, a much-cited study by Mills and Krantz [40] showed that allowing blood donors a choice over which arm to use had a positive effect on donors' experiences in blood transfusion centers. Extending these findings to consumer settings, Hui and Bateson [41] showed that conditions of crowding—generally related to decreases in control as other people in store environments lengthen shopping time and may block access to aisles and products therein—sorted fewer negative effects on control and pleasure for consumers who had a choice to enter a retail setting than for consumers who had no choice.

Translated to the current context, these findings suggest that giving users a choice (via personalization) over how they will be addressed by a Web-based app, or what the app looks like, may increase feelings of control and, thereby, well-being. Moreover, giving people active control may transform otherwise passive patients at the mercy of hostile technology into active citizens responsible for their own well-being.

Personalization at the content level has been used in Web-based interventions only sparsely. An example is the study of Andersson et al [42] where users of a Web-based treatment of anxiety disorders could choose which modules they wanted to engage in.

Figure 4. The Web-based intervention, Na-Aapje (The Netherlands Nutrition Centre).



Design Approach II: Ambient Information

We live in a world of continuous information overflow. A reasonable part of the information we receive is aimed to influence us in some way. Guadagno and Cialdini [43] report how a colleague counted over 500 influence appeals over an hour. Due to the large amount of (influential) information that reaches us during our busy daily routines, we can easily miss out on the information that we *did* want to notice. Not overburdening users and spreading feedback through multiple modalities reduces cognitive workload, rendering interactions with health interventions more pleasurable (P in PERMA) and effective, triggering feelings of accomplishment (A in PERMA). A huge drawback of Web-based interventions in our current information-dense world is that they are for the most part not visible to their users. People may, therefore, easily forget to go online to use the intervention. Moreover, when they do think about going online, they may want to ignore this thought because they did not enjoy working with the intervention before and may not feel the *need* to go online. The design of a Web-based intervention should accommodate for this *need* or *desire*. Designers of Web-based interventions should consider the information overflow that their users will inevitably face, and think carefully about what information they have to give, and at what point or in which context. At the same time, users may need status/progress information that is at hand or in sight in order to evoke a sense of urgency to use the Web-based intervention. Providing information in different places or providing feedback through other modalities than the often-used visual modality could accommodate for this.

XSL•FO
RenderX

Designers have explored both of these approaches in research on the design and effect of ambient information and in explorations on peripheral interaction. The concept of peripheral interaction deals with how technology and interactions that people have with technology is integrated throughout our everyday activities. Peripheral interaction entails that the attention that we pay to technology can at times take place in the periphery of attention. Bakker et al [44] explored how different modalities can be used for peripheral interaction and used both physical (ie, tangible) interaction and auditory displays. The concept of peripheral interaction is similar to approaches that have been used earlier in the design of so-called awareness systems. Such systems have been targeted at displaying one type of information—that of presence of other people at another location—in isolation on a physical object, making this information much more accessible and prominent. An example of such a system is the SnowGlobe [45] that displays movement of a remote user by glowing brighter and that also offers direct opportunities for interaction.

In their work on lifestyle behavior change technologies, Consolvo et al [46] define four design strategies and argue that presenting "abstract" information rather than specific information would have a positive effect on the effectiveness of persuasive systems. This is in line with what Nelson [47] incorporated in his design of Bouncers (see Figure 5). Bouncers is a wallpaper on mobile phones of a group of friends that visualizes everyone's activity through moving circles. In this abstract way, it tells its users about their movements in relation to that of their friends.

As another example of more abstract information display, Ham and Midden [48] compared reactions toward factual (ie, numeric) information and ambient information—light changing color—in the case of information about energy consumption and argue that giving ambient (ie, less specific) information can be more influential because it requires less cognitive capacity to process this information.

In summary, the experience of Web-based interventions can benefit from using different ways to communicate information. Depending on the situation someone is in, and on the importance of the information, designers can choose whether specific or more abstract information is the better choice, and can decide on the specific location of the information (ie, within the Web-based interventions or through a different device or location).

Figure 5. The mobile phone wallpaper, Bouncers, by Terence Nelson.



Design Approach III: Use of Metaphors

One of the key challenges in designing Web-based apps is how to give shape in words and visualizations to challenges, goals, and feedback provided during interaction in order to provide a meaningful and engaging "story" (M and E in PERMA). In many traditional apps, emphasis is on concrete textual input (ie, instructions or numbers signifying scores or tasks left to accomplish) or concrete images (eg, wallpapers or avatars). However, in addition to such concrete elements, much of everyday thought and experience is inherently metaphorical in nature as originally demonstrated by Lakoff and Johnson in their first joint effort, Metaphors We Live By [49]. The central claim therein was that the way we think about abstract concepts,

XSL•FO
RenderX

such as accomplishment, challenge, and perseverance, are metaphorical in nature and are rooted in everyday bodily experiences [49-51]. Of special relevance to the current undertaking was the finding that these embodied metaphors are by no means restricted to language use, but are highly "visual" in nature, as demonstrated by Forceville [52].

For instance, in metaphorical expressions such as "we've made it to the top," "we have a long way to go," and "she was keeping me at a distance," abstract meanings (ie, achievement, lack of communication, psychological support) are talked about in terms of visual-spatial patterns (ie, "rising to the top" or "being close or far away from an end goal"). Such couplings are embodied because they are grounded in everyday bodily interactions in and with our environments.

With respect to object perception, Van Rompay et al [53] showed that verticality or "relative height" is not only the basis of figurative expressions conveying dominance (eg, "He's on top of his game"), but that it also steers meaning perception in design. For instance, in the latter study, everyday products of great vertical size were perceived as more dominant, proud, and impressive compared to products of lesser height.

Furthermore, recent studies have demonstrated that metaphor use or having people behave in line with employed metaphors may also impact feelings and behaviors [50,54,55]. For instance, Gibbs [50] recently showed that a successful relationship metaphor (eg, "Your relationship was moving along in a good direction") prompted people to walk further and longer afterward compared to when they were cued with an unsuccessful relationship metaphor (eg, a metaphor stressing thwarted progress along the road). Another study stressing embodied features of motivation and decision making showed that merely watching a series of squares expanding, versus contracting, prompted participants to feel more self-confident and capable of self-actualization [55].

In design research, various authors have explored metaphor use in relation to affect product experience [56,57]. For instance, Ludden et al [57] showed that product designers can create products incorporating metaphors that people understand by making mind maps revealing common associations from source to target, such that motivating aspects of a source domain (eg, excitement of a journey) can be transferred to the target domain (ie, the product).

Taken together, these examples and insights suggest that employing metaphors in Web-based apps may not just be fun or trivial, but that it may actually contribute to adherence in the long run by creating meaning and fostering engagement. Indeed, metaphors have been used for the design of mobile apps that seek to motivate people toward certain types of behavior. A well-known example in literature on persuasive design is the design of a flourishing garden that was used by Consolvo et al [58] as a metaphor for a flourishing (ie, active and satisfying) life.

In summary, depending on the type of app and the goals set out for the app, different types of metaphorical mappings may be drawn upon that may become ever more influential as we move from static, to dynamic, to fully interactive apps in the years to come. Not only can such visual metaphors enhance fun and engagement (positive emotions, P), they may also render goals that people set out for, and make the activities they undertake to achieve them, more meaningful and worthwhile, precisely because they connect to people's intuitive understanding of the world around them.

A nice example of an endeavor to use a metaphor in a Web-based intervention context is the intervention, This Is Your Life!, that was based on an existing positive psychology intervention [59], and designed for a specific target group: teachers at a primary school (see Ludden et al [60] for a full report of the design process). The designers of the interactive part of the intervention worked in close collaboration with the psychologists who designed the content of the intervention, and in close interaction with the proposed target group. A set of three concept designs was made in which different types of source domains (ie, a library, tree, and journey, respectively) were used (see Figure 6).

The concept of the intervention presented as a journey on a map—favored by all 8 participants in a focus group—was further developed into a working prototype (see Figure 7). For the final prototype, the metaphor of the journey was further developed into details of the intervention. For example, the typical terminology from training or school-like activities (ie, lessons, exercises, chapters) was changed into terminology that was more relevant to the metaphor of the journey. For instance, *chapters* are *locations* on the map and each location has *challenges* which are the *exercises* of the training. When a user completes the challenges for a specific location, he or she can get the "key" to the next location. Locations also have names that are related to the "life is a journey" metaphor, such as "The island of broken dreams" and "The river of flow."

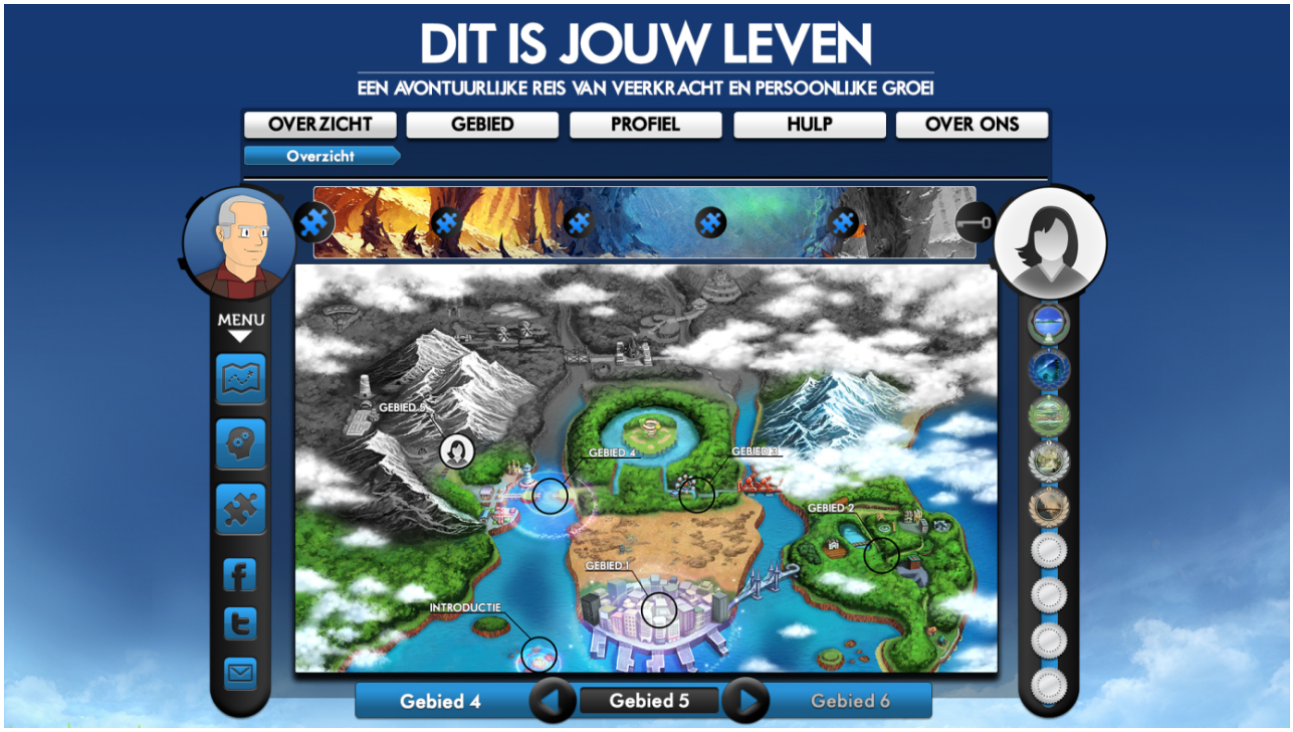
From short, individual, evaluative sessions with primary school teachers, the developers of this intervention found that the metaphor that was used for the design of the intervention was seen as motivating and stimulating by users.

Figure 6. Three concept designs from the Web-based intervention, This Is Your Life! Concepts: library, tree of happiness, and journey on a map.



XSL•FO
RenderX

Figure 7. Final prototype of the Web-based intervention, This Is Your Life!



Discussion

The Role of Design(ers)

We have argued that designers of technology for well-being could stimulate well-being in two ways: by designing a positive experience during use and by developing persuasive effective interventions that positively influence adherence, leading to a higher level of overall well-being for the user. To specify how design can contribute, we have introduced three approaches that have recently gained particular attention in design research. In doing so, we have tried to broaden the discussion on the role of design in the development of Web-based interventions and eHealth in general. So far, this discussion has been mainly focused on the need to involve users and usability experts in the design process [61,62], as well as on transferring knowledge from behavioral sciences about persuasion techniques [63]. This has in most cases lead to detailed "design guidelines" about the organization and specific features of eHealth systems, factors that are, of course, of influence on how the interventions will be experienced. However, design guidelines that are more focused on the different ways in which these factors can eventually take shape and how this can have an effect on the experience of the user have so far mainly been neglected. That said, next to the three approaches introduced here, various other design approaches could also contribute to one of the elements of PERMA. We have chosen the three discussed here because they seem to be prevalent choices in recent design research in the eHealth domain. Inspiration for other approaches could, for example, be found in literature on how to apply elements of gaming in a therapy context [64]. However, the best way to ensure that all aspects contribute to a desired experience for an eHealth intervention is to facilitate cooperation between designers of the content of this technology (eg, the therapists developing health interventions), designers of the system (eg, human-computer interaction [HCI] developers), and designers of the form of the technology (eg, interaction designers or product designers) from an early stage of development, and discuss which design features and organization can enrich the presentation of the content. Together, this team of designers can reflect on questions such as “What seem to be the most promising new approaches/theories/tools?” “How will we apply these to practice?” “What is the future of Web-based health interventions?” and “How to incorporate more interactive and persuasive features?” This last question is also reflected in one of the principles of the holistic framework to improve the uptake and impact of eHealth technologies as described by van Gemert-Pijnen et al [64]. As we have argued and illustrated throughout this paper, involving users at an early stage of the design process can certainly be a valuable way to make informed design decisions on which persuasive/motivational elements to use.

Future Research

In this paper, three approaches that have gained particular attention in design research were introduced and their possibilities to serve the design of Web-based interventions were discussed. Clearly, follow-up research is needed to further pinpoint the shortcomings in the current design of Web-based interventions, to test whether the approaches discussed indeed are effective, and to suggest how they may be applied in ever more advanced and interactive technologies.

For example, application possibilities of the design tools discussed here may also lie in third generation eHealth interventions. In this context, design can play an important role in, for example, the way feedback on behavior is offered to people through mobile and/or wearable coaching systems. Next to this, as Web-based apps or games are becoming more interactive—think, for instance, of serious gaming apps making

XSL•FO
RenderX

use of a Wii or PlayStation Move controller—means for enriched metaphor portrayal are growing all the time. Hence, in addition to watching expanding stimuli (eg, expanding circles on a water surface) and feeling more self-confident—compare with Ludden et al [57]—such interactive game controllers may also prompt people to take in specific postures. Hence, when cultivation of open-mindedness is at stake, people may be prompted to take in an expansive, as opposed to a contracted, posture. This is again in line with linguistic expressions in which openness and intellectual growth are coupled, for example, "open up to others," "that blew my mind," and "her world is so small." Interestingly, the latter relationship between taking in an expansive posture and feeling more confident was recently demonstrated by Carney et al [54]. Specifically, they showed that people taking in an expansive bodily posture displayed more self-confidence and were less likely to seek compromise in a negotiation task. Hence, future research could address means by which "priming" characteristics, such as open-mindedness and self-confidence, may contribute to adherence in the long run.

As for reach and tailored communication with specific target groups, it is important to keep in mind that different target groups may react differently depending on metaphor choice and personalization options. Recently, Halko and Kientz [35] showed that the type of appeal exerted by persuasive technologies may indeed sort different effects dependent on target group personalities, with a more authoritarian style of interaction, for instance, faring better with some target groups than others.

As for personalization and feelings of control, numerous studies have likewise shown that people vary in the extent to which they seek and appreciate control over events in their lives [65,66]. People high in desire for control generally seek influence over others and desire control over events, and hence react with a strong (negative) emotional response when opportunities for control are lacking. In the online context, this may thus lead to nonadherence and dropout. People low in desire for control, on the other hand, are less likely to react with negative affect when faced with a perceived inability to control events, and may appreciate it when decisions are made by others.

Prior to metaphor selections and making decisions on how specific feedback information should be, on whether to incorporate nonvisual information, and on whether or not to incorporate personalization options in Web apps, insight into the values and needs of different target groups is a prerequisite. In the online context, such insights may come in the form of short intake questionnaires, based on which such decisions can be made. In conclusion, using design research as a source of inspiration for the design of Web-based interventions, as outlined in this paper, is a promising means to optimize the reach of, and adherence to, such interventions. Moreover, cooperation between different types of designers of interventions is, next to involving the envisioned users of an intervention, essential to come to a design that incorporates all aspects (ie, content, form, organization) that can contribute to a desired experience for the end user. Moreover, this cooperation should start in early phases of the development, not after a "who, what, and why" has been proposed for an intervention. Early involvement of all types of designers will allow for more innovative designs in which the different elements of the intervention can be experienced in an integrated way.

Conflicts of Interest

None declared.

References

1. Diener E, Chan MY. Happy people live longer: subjective well-being contributes to health and longevity. Appl Psychol Health Well Being 2011 Mar;3(1):1-43. [doi: 10.1111/j.1758-0854.2010.01045.x]
2. Dolan P, Peasgood T, White M. Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being. J Econ Psychol 2008 Feb;29(1):94-122.
3. Howell RT, Kern ML, Lyubomirsky S. Health benefits: meta-analytically determining the impact of well-being on objective health outcomes. Health Psychol Rev 2007;1(1):83-136.
4. Steptoe A, Wardle J, Marmot M. Positive affect and health-related neuroendocrine, cardiovascular, and inflammatory processes. Proc Natl Acad Sci U S A 2005 May 3;102(18):6508-6512 [FREE Full text] [doi: 10.1073/pnas.0409174102] [Medline: 15840727]
5. Lamers SM, Bolier L, Westerhof G, Smit F, Bohlmeijer ET. The impact of emotional well-being on long-term recovery and survival in physical illness: a meta-analysis. J Behav Med 2012 Oct;35(5):538-547 [FREE Full text] [doi: 10.1007/s10865-011-9379-8] [Medline: 21918889]
6. Seligman MEP. Flourish: A Visionary New Understanding of Happiness and Well-being. New York, NY: The Free Press; 2011.
7. Cappuccio F, Kerry S, Forbes L, Donald A. Blood pressure control by home monitoring: meta-analysis of randomised trials. BMJ 2004 Jul 17;329(7458):145 [FREE Full text] [doi: 10.1136/bmj.38121.684410.AE] [Medline: 15194600]
8. Barak A, Hen L, Boniel-Nissim M, Shapira NA. A comprehensive review and a meta-analysis of the effectiveness of Internet-based psychotherapeutic interventions. J Technol Hum Serv 2008;26(2-4):109-160.
9. Chittaro L, Vianello A. Computer-supported mindfulness: evaluation of a mobile thought distancing application on naïve meditators. Int J Hum Comput Stud 2014 Mar;72(3):337-348.

XSL•FO
RenderX

10. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to Web-based interventions. J Med Internet Res 2012;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
11. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]
12. Postel M, de Jong CA, de Haan HA. Does e-therapy for problem drinking reach hidden populations? Am J Psychiatry 2005 Dec;162(12):2393. [doi: 10.1176/appi.ajp.162.12.2393] [Medline: 16330613]
13. Rothert K, Strecher V, Doyle L, Caplan W, Joyce J, Jimison H, et al. Web-based weight management programs in an integrated health care setting: a randomized, controlled trial. Obesity 2006 Feb;14(2):266-272. [doi: 10.1038/oby.2006.34] [Medline: 16571852]
14. Binks M, van Mierlo T. Utilization patterns and user characteristics of an ad libitum Internet weight loss program. J Med Internet Res 2010;12(1):e9 [FREE Full text] [doi: 10.2196/jmir.1347] [Medline: 20350926]
15. 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]
16. Kelders SM, Bohlmeijer ET, Van Gemert-Pijnen JE. Participants, usage, and use patterns of a Web-based intervention for the prevention of depression within a randomized controlled trial. J Med Internet Res 2013;15(8):e172 [FREE Full text] [doi: 10.2196/jmir.2258] [Medline: 23963284]
17. Balmford J, Borland R, Benda P. Patterns of use of an automated interactive personalized coaching program for smoking cessation. J Med Internet Res 2008;10(5):e54 [FREE Full text] [doi: 10.2196/jmir.1016] [Medline: 19097975]
18. Norman GJ, Zabinski MF, Adams MA, Rosenberg DE, Yaroch AL, Atienza AA. A review of eHealth interventions for physical activity and dietary behavior change. Am J Prev Med 2007 Oct;33(4):336-345 [FREE Full text] [doi: 10.1016/j.amepre.2007.05.007] [Medline: 17888860]
19. IJsselsteijn W, de Kort Y, Midden C, Eggen B, van den Hoven E. Persuasive technology for human well-being: setting the scene. In: IJsselsteijn W, de Kort Y, Midden C, Eggen B, van den Hoven E, editors. Persuasive Technology. Berlin, Germany: Springer Berlin Heidelberg; 2006:1-5.
20. Tromp N, Hekkert P, Verbeek P. Design for socially responsible behavior: a classification of influence based on intended user experience. Design Issues 2011 Jul;27(3):3-19. [doi: 10.1162/DESI_a_00087]
21. Pohlmeyer A. Design for happiness. Interfaces 2012;92:8-11.
22. Postel M, de Haan HA, Ter Huurne ED, Becker ES, de Jong CA. Characteristics of problem drinkers in e-therapy versus face-to-face treatment. Am J Drug Alcohol Abuse 2011 Nov;37(6):537-542. [doi: 10.3109/00952990.2011.600388] [Medline: 21797813]
23. 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]
24. Linke S, Brown A, Wallace P. Down your drink: a Web-based intervention for people with excessive alcohol consumption. Alcohol Alcohol 2004;39(1):29-32 [FREE Full text] [Medline: 14691071]
25. Linke S, McCambridge J, Khadjesari Z, Wallace P, Murray E. Development of a psychologically enhanced interactive online intervention for hazardous drinking. Alcohol Alcohol 2008;43(6):669-674 [FREE Full text] [doi: 10.1093/alcalc/agn066] [Medline: 18693217]
26. Eysenbach G. Consumer health informatics. BMJ 2000 Jun 24;320(7251):1713-1716 [FREE Full text] [Medline: 10864552]
27. Hart JT. The inverse care law. Lancet 1971 Feb 27;1(7696):405-412. [Medline: 4100731]
28. Schubart JR, Stuckey HL, Ganeshamoorthy A, Sciamanna CN. Chronic health conditions and internet behavioral interventions: a review of factors to enhance user engagement. Comput Inform Nurs 2011 Feb;29(2):81-92. [doi: 10.1097/NCN.0b013e3182065eed] [Medline: 21164337]
29. Brouwer W, Kroeze W, Crutzen R, de Nooijer J, de Vries NK, Brug J, et al. Which intervention characteristics are related to more exposure to Internet-delivered healthy lifestyle promotion interventions? A systematic review. J Med Internet Res 2011;13(1):e2 [FREE Full text] [doi: 10.2196/jmir.1639] [Medline: 21212045]
30. Oinas-Kukkonen H, Harjumaa M. Persuasive Systems Design: key issues, process model, and system features. Communications of the Association for Information Systems 2009 Mar;24(1):485-500.
31. Titov N, Andrews G, Choi I, Schwencke G, Johnston L. Randomized controlled trial of Web-based treatment of social phobia without clinician guidance. Aust N Z J Psychiatry 2009 Jan;43(10):913-919. [doi: 10.1080/00048670903179160]
32. Woodall WG, Buller DB, Saba L, Zimmerman D, Waters E, Hines JM, et al. Effect of emailed messages on return use of a nutrition education website and subsequent changes in dietary behavior. J Med Internet Res 2007;9(3):e27 [FREE Full text] [doi: 10.2196/jmir.9.3.e27] [Medline: 17942389]
33. Fogg BJ. Persuasive Technology: Using Computers to Change What We Think and Do. San Francisco, CA: Morgan Kaufmann Publishers; 2003.
34. Seligman MEP. Authentic Happiness: Using the New Positive Psychology to Realize Your Potential for Lasting Fulfillment. New York, NY: The Free Press; 2002.

XSL•FO
RenderX

35. Halko S, Kientz J. Personality and persuasive technology: an exploratory study on health-promoting mobile applications. In: Ploug T, Hasle P, Oinas-Kukkonen H, editors. Persuasive Technology. Berlin, Germany: Springer Berlin Heidelberg; 2010:150-161.
36. Kaptein M, De Ruyter B, Markopoulos P, Aarts E. Adaptive persuasive systems: a study of tailored persuasive text messages to reduce snacking. ACM Trans Interact Intell Syst 2012 Jun 01;2(2):1-25. [doi: 10.1145/2209310.2209313]
37. Lee MK, Kiesler S, Forlizzi J. Mining behavioral economics to design persuasive technology for healthy choices. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2011 Presented at: SIGCHI Conference on Human Factors in Computing Systems; May 7-12, 2011; Vancouver, BC p. 325-334.
38. Mehrabian A, Russel JA. An Approach to Environmental Psychology. Cambridge, MA: MIT Press; 1974.
39. Averill JR. Personal control over aversive stimuli and its relationship to stress. Psychol Bull 1973 Oct;80(4):286-303.
40. Mills RT, Krantz DS. Information, choice, and reactions to stress: a field experiment in a blood bank with laboratory analogue. J Pers Soc Psychol 1979 Apr;37(4):608-620. [Medline: 376826]
41. Hui MK, Bateson JEG. Perceived control and the effects of crowding and consumer choice on the service experience. J Consum Res 1991 Sep;18(2):174-184.
42. Andersson G, Estling F, Jakobsson E, Cuijpers P, Carlbring P. Can the patient decide which modules to endorse? An open trial of tailored Internet treatment of anxiety disorders. Cogn Behav Ther 2011;40(1):57-64. [doi: 10.1080/16506073.2010.529457] [Medline: 21337215]
43. Guadagno R, Cialdini RB. Online persuasion: an examination of gender differences in computer-mediated interpersonal influence. Group Dyn 2002;6(1):38-51. [doi: 10.1037/1089-2699.6.1.38]
44. Bakker S, van den Hoven E, Eggen B. Peripheral interaction: characteristics and considerations. Pers Ubiquitous Comput 2015 Jan;19(1):239-254. [doi: 10.1007/s00779-014-0775-2]
45. Visser T, Vastenburg MH, Keyson DV. Designing to support social connectedness: the case of SnowGlobe. International Journal of Design 2011;5(3):129-142.
46. Consolvo S, McDonald D, Landay J. Theory-driven design strategies for technologies that support behavior change in everyday life. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2009 Presented at: SIGCHI Conference on Human Factors in Computing Systems; April 4-9, 2009; Boston, MA p. 405-414.
47. Nelson TBM, Megens CJPG, Peeters MMR. Bouncers: a design exploration into sharing and triggering personal activity amongst friends. In: Proceedings of the 7th International Conference on Persuasive Technology. Sweden: Springer; 2012 Presented at: The 7th International Conference on Persuasive Technology; June 6-8, 2012; Linköping, Sweden p. 14-18.
48. Ham J, Midden C. Ambient persuasive technology needs little cognitive effort: the differential effects of cognitive load on lighting feedback versus factual feedback. In: Ploug T, Hasle P, Oinas-Kukkonen H, editors. Persuasive Technology. Berlin, Germany: Springer Berlin Heidelberg; 2010:132-142.
49. Lakoff G, Johnson M. Metaphors We Live By. Chicago, IL: The University of Chicago Press; 1980.
50. Gibbs R. Walking the walk while thinking about the talk: embodied interpretation of metaphorical narratives. J Psycholinguist Res 2013 Aug;42(4):363-378. [doi: 10.1007/s10936-012-9222-6] [Medline: 22585389]
51. Johnson M. The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason. Chicago, IL: The University of Chicago Press; 1987.
52. Forceville C. Pictorial Metaphor in Advertising. Abingdon, UK: Routledge; 1996.
53. van Rompay T, Hekkert P, Saakes D, Russo B. Grounding abstract object characteristics in embodied interactions. Acta Psychol (Amst) 2005 Jul;119(3):315-351. [doi: 10.1016/j.actpsy.2005.02.001] [Medline: 15939029]
54. Carney DR, Cuddy AJ, Yap AJ. Power posing: brief nonverbal displays affect neuroendocrine levels and risk tolerance. Psychol Sci 2010 Oct;21(10):1363-1368. [doi: 10.1177/0956797610383437] [Medline: 20855902]
55. Landau MJ, Vess M, Arndt J, Rothschild ZK, Sullivan D, Atchley RA. Embodied metaphor and the “true” self: priming entity expansion and protection influences intrinsic self-expressions in self-perceptions and interpersonal behavior. J Exp Soc Psychol 2011 Jan;47(1):79-87. [doi: 10.1016/j.jesp.2010.08.012]
56. Cila N. Metaphors We Design By: The Use of Metaphors in Product Design [dissertation]. Delft, The Netherlands: Delft University of Technology; Sep 26, 2013.
57. Ludden GDS, Kudrowitz BM, Schifferstein HNJ, Hekkert P. Surprise and humor in product design. Designing sensory metaphors in multiple modalities. Humor 2012;25(3):285-309. [doi: 10.1515/humor-2012-0015]
58. Consolvo S, McDonald DW, Toscos T, Chen MY, Froehlich J, Harrison B, et al. Activity sensing in the wild: a field trial of ubifit garden. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2008 Presented at: SIGCHI Conference on Human Factors in Computing Systems; April 5-10, 2008; Florence, Italy p. 1797-1806.
59. Bohlmeijer ET, Hulsbergen ML. Dit Is Jouw Leven: Ervaar de Effecten van de Positieve Psychologie. Amsterdam, The Netherlands: Uitgeverij Boom; 2013.
60. Ludden GDS, Kelders SM, Snippert BHJ. 'This Is Your Life!': the design of a positive psychology intervention using metaphor to motivate. In: Spagnolli A, Chittaro L, Gamberini L, editors. Persuasive Technology. Cham, Switzerland: Springer International Publishing; 2014:179-190.
61. Doherty G, Coyle D, Matthews M. Design and evaluation guidelines for mental health technologies. Interact Comput 2010 Jul;22(4):243-252. [doi: 10.1016/j.intcom.2010.02.006]

XSL•FO
RenderX

62. Lederman R, Wadley G, Gleeson J, Bendall S, Álvarez-Jiménez M. Moderated online social therapy. ACM Trans Comput Hum Interact 2014 Feb 01;21(1):1-26. [doi: 10.1145/2513179]
63. Visch VT, de Wit M, Dinh L, vandenBrule D, Melles M, Sonneveld MH. Industrial design meets mental healthcare: designing therapy-enhancing products involving game-elements for mental healthcare - three case studies. In: IEEE 1st International Conference on Serious Games and Applications for Health (SeGAH). 2011 Presented at: IEEE Serious Games and Applications for Health (SEGAH); November 16-18, 2011; Braga, Portugal p. 1-6.
64. van Gemert-Pijnen JE, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
65. Burger JM, Cooper HM. The desirability of control. Motiv Emot 1979 Dec;3(4):381-393. [doi: 10.1007/BF00994052]
66. Gebhardt WA, Brosschot JF. Desirability of control: psychometric properties and relationships with locus of control, personality, coping, and mental and somatic complaints in three Dutch samples. Eur J Pers 2002 Nov;16(6):423-438. [doi: 10.1002/per.463]

Abbreviations

COPD: chronic obstructive pulmonary disease
HCI: human-computer interaction
PERMA: positive emotions, engagement, positive relationships, meaning, and accomplishment

Edited by G Eysenbach; submitted 05.01.15; peer-reviewed by A Pohlmeyer, P Tikka; comments to author 19.03.15; revised version received 02.05.15; accepted 24.05.15; published 10.07.15.

Please cite as:
Ludden GDS, van Rompay TJL, Kelders SM, van Gemert-Pijnen JEWC
How to Increase Reach and Adherence of Web-Based Interventions: A Design Research Viewpoint
J Med Internet Res 2015;17(7):e172
URL: http://www.jmir.org/2015/7/e172/
doi:10.2196/jmir.4201
PMID:26163456

©Geke DS Ludden, Thomas JL van Rompay, Saskia M Kelders, Julia EWC van Gemert-Pijnen. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 10.07.2015. 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 Web-Based, Social Networking Physical Activity Intervention for Insufficiently Active Adults Delivered via Facebook App: Randomized Controlled Trial

Carol Maher[1], PhD; Monika Ferguson[1], PhD; Corneel Vandelanotte[2], PhD; Ron Plotnikoff[3], PhD; Ilse De Bourdeaudhuij[4], PhD; Samantha Thomas[5], PhD; Karen Nelson-Field[6], PhD; Tim Olds[1], PhD

[1]Alliance for Research in Exercise, Nutrition and Activity, University of South Australia, Adelaide, Australia

[2]Institute for Health and Social Sciences, Central Queensland University, Adelaide, Australia

[3]Priority Research Centre in Physical Activity and Nutrition, University of Newcastle, Newcastle, Australia

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

[5]School of Health and Society, University of Wollongong, Wollongong, Australia

[6]Ehrenberg-Bass Institute for Marketing Science, University of South Australia, Adelaide, Australia

Corresponding Author:
Carol Maher, PhD
Alliance for Research in Exercise, Nutrition and Activity
University of South Australia
GPO Box 2471
Adelaide, 5001
Australia
Phone: 61 08 830 ext 22315
Fax: 61 8 8302 2315
Email: carol.maher@unisa.edu.au

Abstract

Background: Online social networks offer considerable potential for delivery of socially influential health behavior change interventions.

Objective: To determine the efficacy, engagement, and feasibility of an online social networking physical activity intervention with pedometers delivered via Facebook app.

Methods: A total of 110 adults with a mean age of 35.6 years (SD 12.4) were recruited online in teams of 3 to 8 friends. Teams were randomly allocated to receive access to a 50-day online social networking physical activity intervention which included self-monitoring, social elements, and pedometers (“Active Team” Facebook app; n=51 individuals, 12 teams) or a wait-listed control condition (n=59 individuals, 13 teams). Assessments were undertaken online at baseline, 8 weeks, and 20 weeks. The primary outcome measure was self-reported weekly moderate-to-vigorous physical activity (MVPA). Secondary outcomes were weekly walking, vigorous physical activity time, moderate physical activity time, overall quality of life, and mental health quality of life. Analyses were undertaken using random-effects mixed modeling, accounting for potential clustering at the team level. Usage statistics were reported descriptively to determine engagement and feasibility.

Results: At the 8-week follow-up, the intervention participants had significantly increased their total weekly MVPA by 135 minutes relative to the control group (P=.03), due primarily to increases in walking time (155 min/week increase relative to controls, P<.001). However, statistical differences between groups for total weekly MVPA and walking time were lost at the 20-week follow-up. There were no significant changes in vigorous physical activity, nor overall quality of life or mental health quality of life at either time point. High levels of engagement with the intervention, and particularly the self-monitoring features, were observed.

Conclusions: An online, social networking physical activity intervention with pedometers can produce sizable short-term physical activity changes. Future work is needed to determine how to maintain behavior change in the longer term, how to reach at-need populations, and how to disseminate such interventions on a mass scale.

XSL•FO
RenderX

Trial Registration: Australian New Zealand Clinical Trials Registry (ANZCTR): ACTRN12614000488606; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=366239 (Archived by WebCite at http://www.webcitation.org/6ZVtu6TMz).

(J Med Internet Res 2015;17(7):e174) doi:10.2196/jmir.4086

KEYWORDS

social network; behavior change; intervention; Internet; physical activity

Introduction

Physical inactivity is a leading modifiable cause of death and disease worldwide and causes as many deaths as smoking [1]. Just 30 minutes a day of moderate-intensity physical activity halves the risk of leading causes of morbidity and mortality, such as cardiovascular disease, type 2 diabetes, and obesity, and reduces the risk of breast and bowel cancer, depression, and anxiety [2]. Despite this, many people in developed countries are insufficiently active to achieve these benefits. For example, a recent nationally representative survey of 20,426 Australians found that 67% of adults self-reported that they got less than 30 minutes of physical activity a day [3]. Population-based interventions are needed to assist the general adult population to become more physically active.

Web-based physical activity interventions offer an opportunity to reach a large number of people at a relatively low cost. Systematic reviews and meta-analyses of Web-based physical activity interventions demonstrate they are effective in changing behavior [4,5], however, typically they have not been adopted by large numbers of users and appear to have difficulty sustaining user engagement over an extended period [4,5]. New intervention approaches capitalizing on recent technology trends, such as online social networks [6] and gamification, may assist in overcoming these issues. Our recent review of online social networks for delivery of health behavior interventions found fledgling, but promising, evidence of effectiveness [7].

Online social networks reportedly account for one-quarter of all time spent online [8,9], and appear to offer considerable potential for delivery of public health campaigns for several reasons. Like the Internet in general, they can reach very large audiences (eg, Facebook, the world's largest social networking website, had 1.32 billion users each month as of June 2014 [10]). They also offer some key advantages over conventional online delivery, including that messages can be delivered via existing social contacts, which may be more influential than health messages delivered via traditional marketing strategies [11]. Furthermore, unlike traditional Web-based interventions [4], online social networks typically achieve high levels of user engagement and retention [12].

Another online trend that has emerged in recent years is gamification. Gamification refers to the application of video game elements, such as fun, challenges, competition, and rewards, in nongaming situations [13]. In the commercial sector, such techniques have reported to markedly increase engagement (eg, a software company reported an 8-fold increase in user engagement after introduction of gamification features) [14]. A recent systematic review of health behavior change interventions delivered using online social networks [7] found that compared to studies which did not incorporate gamification features, the one study that did—in the form of competition between users [15]—achieved substantially larger intervention effects and higher levels of user engagement.

To date, only a handful of studies have attempted to use existing popular online social network platforms, such as Facebook and Twitter, to intervene on physical activity. The most common approach has been the use of a Twitter feed or private Facebook groups to share content regarding physical activity and facilitate discussion between study participants [7]. In most cases, the online social network intervention has been provided as a component, complementing a more comprehensive intervention package, for example, involving access to a physical activity self-monitoring website with personalized feedback from a health professional [16]; provision of pedometer, digital scales, cooking equipment, and personalized feedback [17]; or access to a series of podcasts, advice from an expert moderator, and a calorie-counting app [18]. To our knowledge, only one previous study [15] has utilized a Facebook app (ie, software created by third party developers to function within the Facebook platform and access data in Facebook) to intervene on physical activity.

The primary objective of this study was to determine whether a team-based 50-day social networking physical activity intervention delivered via Facebook app and incorporating gamification features was effective in changing weekly moderate-to-vigorous physical activity (MVPA) in adults aged 18 to 65 years. The secondary objectives were (1) to determine whether the intervention impacted other physical activity (ie, weekly walking, vigorous physical activity, and moderate physical activity time) and quality of life (in particular, mental quality of life), (2) to determine usage, and (3) to examine the feasibility of the online intervention.

Methods

Overview

Ethical approval for this cluster randomized controlled trial (RCT) was obtained from the University of South Australia Human Research Ethics Committee, and the study was registered with the Australian and New Zealand Clinical Trials Registry, protocol number: ACTRN12614000488606. Data collection took place between September 2013 and July 2014. Participants provided informed consent online prior to commencing the study. The study was designed, and the manuscript prepared, following CONSORT guidelines [19].

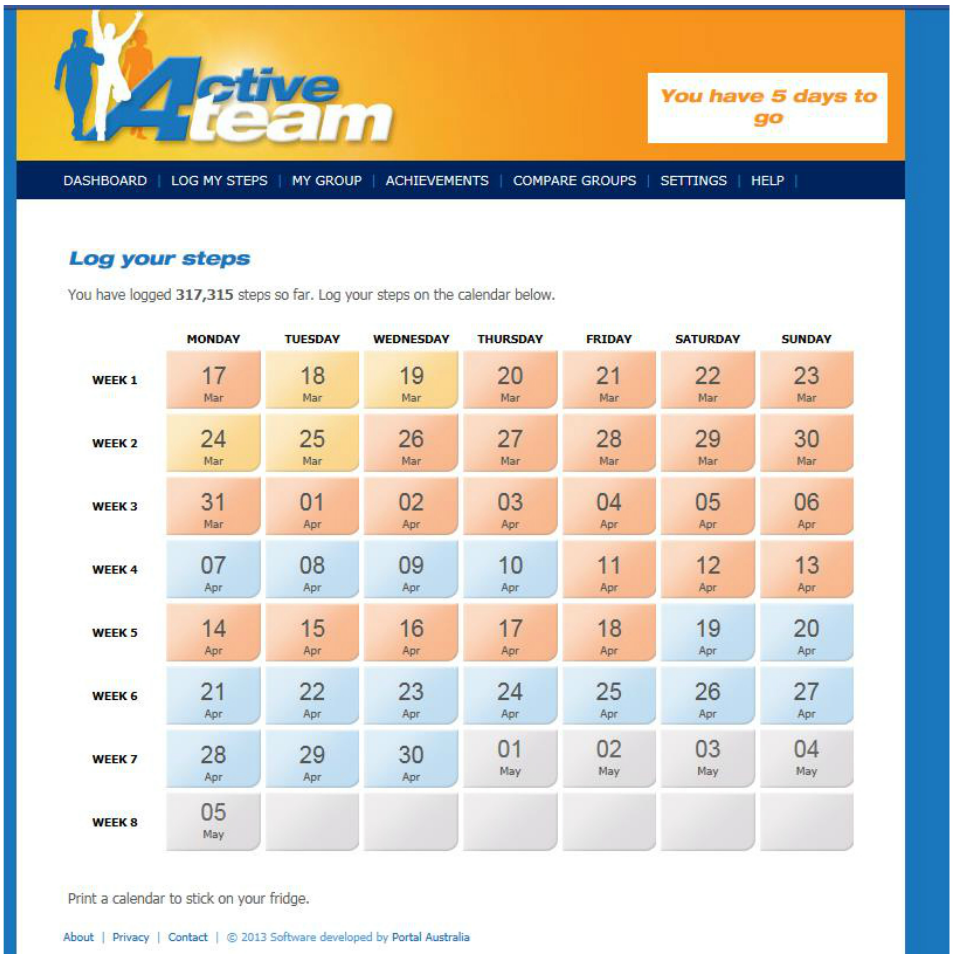
XSL•FO
RenderX

Intervention

Active Team is a new, free, 50-day team-based Facebook app, developed to assist adults to increase their physical activity levels. The content and features of the program were developed by a team at the University of South Australia led by Dr Carol Maher, following a series of interviews with 20 adults regarding the potential for developing a physical activity intervention delivered via online social networks (unpublished). Commercial software developers were engaged to produce the software platform, and extensive pilot-testing and usability testing was undertaken for the first version of the software [20]. Participants are provided with a pedometer, and encouraged to achieve 10,000 steps per day [21], working in teams of 3 to 8 existing Facebook friends. Active Team is designed to encourage friendly rivalry within friendship groups, offer peer encouragement and support, and be quick, social, and enjoyable to use. It includes a calendar to log daily step counts (steps can be logged up to 7 days in arrears) (see Figure 1); a dashboard showing step-logging progress, awards, and gifts (see Figure 2); a team tally board to allow users to monitor their own and their teammates' progress; a team message board for team members to communicate with one another; daily tips for increasing physical activity; gamification features, such as awards for individual and team step-logging and step-count achievements; and the ability to send virtual gifts to teammates. Automated computer-tailored weekly emails are sent to participants summarizing their progress and encouraging continued participation. Apart from provision of a pedometer, the Active Team intervention approach was designed to be minimally resource intensive and, therefore, did not include provision of extensive instrumental support, expert moderation, or feedback from a health professional.

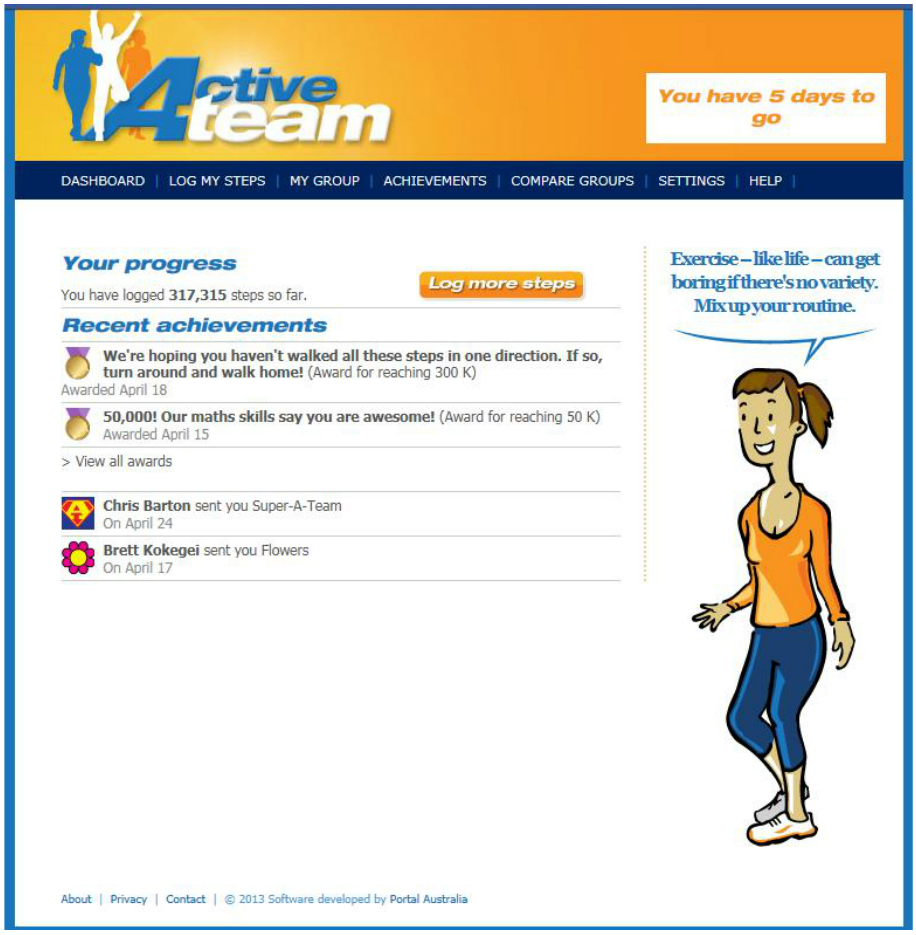
Following consideration of numerous behavior change theories, the theory of planned behavior [22,23] and fun theory [24] were selected to inform development of the content and features of Active Team. The theory of planned behavior posits that a person's decision to perform a particular behavior is influenced by three factors: attitude, subjective norms, and perceived behavioral control [22,23]. Fun theory advocates that people will be more motivated to do routine activities if they are adapted to be fun [24]. The Active Team app attempts to address each of these factors by providing daily tips for physical activity, written by a comedian (theory of planned behavior—attitudes and perceived behavioral control; fun theory); use of teams for peer encouragement and support (theory of planned behavior—subjective norms; fun theory); and setting small achievable goals (ie, daily step count), which are recorded and contribute to a long-term/overall goal (500,000 steps) (theory of planned behavior—attitude and perceived behavioral control), unlockable awards, named by a comedian (fun theory), and the ability to send virtual gifts, such as a high five and a pink leotard (theory of planned behavior—subjective norms; fun theory).

Figure 1. Active Team step-logging calendar.



XSL•FO
RenderX

Figure 2. Active Team dashboard, showing step-logging progress, awards, and gifts.



Participants and Procedures

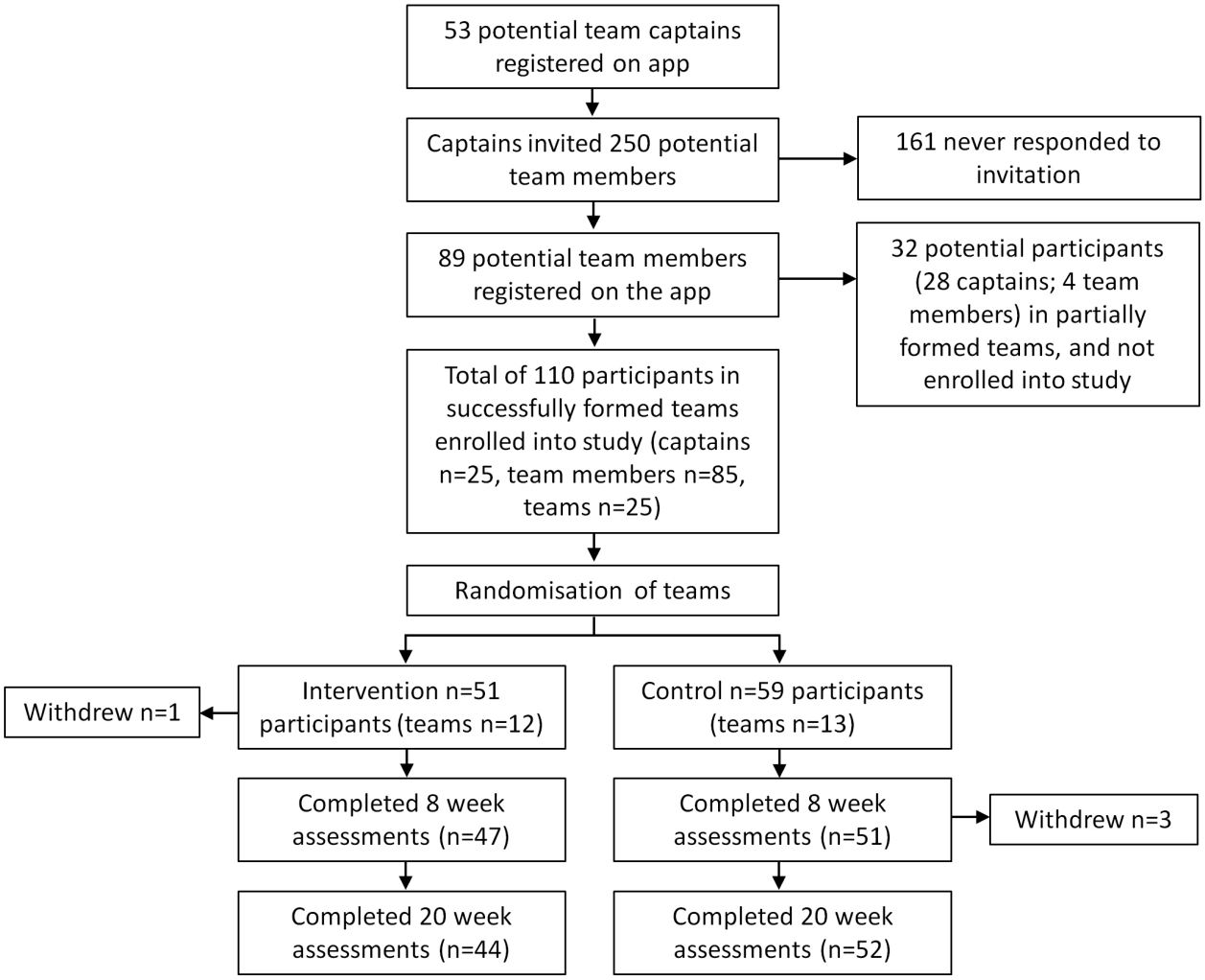
An overview of the randomized controlled trial methodology is shown in Figure 3. Participants were recruited through a Facebook advertising campaign, media stories in the local newspaper and television news bulletin, and distribution of flyers at the University of South Australia campuses. Participants were eligible if they met the following criteria: (1) were between the ages of 18 and 65 years, (2) considered themselves insufficiently active (ie, not currently achieving the Australian guidelines of 150 min of MVPA/week), (3) were current Facebook users, (4) did not have an existing medical condition for which they had been advised by a doctor to avoid exercise, and (5) were able to speak English.

Interested participants could access the app by typing “Active Team” into the Facebook search function, or by following a link included in the Facebook advertisement. The first page of the app was a welcome page, containing an information video and a detailed participant information sheet. Participants could then use the app to register interest in the study and complete baseline surveys. The app guided participants through the process of inviting eligible Facebook friends to form a team, which resulted in an invitation being posted on relevant friends’ Facebook newsfeeds, along with a link to the app. Participants were formally enrolled into the study if they completed baseline surveys and were part of a team comprising 3 to 8 members. Once a team was finalized, the whole team was randomly allocated to either the intervention or the control condition, using a computer-generated randomization sequence with blocking (block size = six) with allocation concealment. Participants received an automated email informing them of which condition they were enrolled in and when their Active Team challenge would begin.

Teams allocated to the intervention condition received access to the full Active Team app and were mailed a pedometer. Teams allocated to the control condition were placed on a waiting list to receive access to the intervention (app and pedometer) at completion of the study and were told that their health would be monitored over the ensuing 5 months.

XSL•FO
RenderX

Figure 3. Overview of participant recruitment, assessment, and flow.



Assessments and Outcome Measures

There were three assessment points for all participants: (1) baseline (at recruitment), (2) 8 weeks (coinciding with the final week of the intervention), and (3) 20 weeks (3-months postcompletion of the intervention). All measurements were completed online. Blinding of participants was not possible, however single blinding was achieved, in that the outcome measures were administered by computer. Thus, the potential for introducing bias was eliminated compared with a scenario where the outcome measures were administered by a person who was aware of the participants' group allocation.

The primary outcome measure was self-reported total weekly MVPA. This was assessed using the Active Australia Survey (AAS) [25], which records physical activities over the previous 7 days. The validated instrument includes eight items relating to the frequency (four items) and duration (four items) of the following: walking (for exercise, recreation, or transport), vigorous physical activities (such as jogging, cycling, aerobics, and competitive sport), and moderate physical activity (such as gentle swimming, tennis, and golf; excluding walking). As per AAS protocol, total weekly MVPA was calculated as walking time + moderate time + (2 x vigorous time), with each individual item being truncated at a maximum of 840 minutes per week, and total physical activity (PA) being truncated at a maximum of 1680 minutes per week in order to reduce the risk of overreporting [25]. The AAS has been shown to have moderate reliability (rho=.56 to .64) [26] and moderate validity when compared with weekly pedometer step counts (rho = .43) and accelerometry (rho =.52) [26].

Secondary outcomes included examining the physical activity types/intensities separately (ie, weekly walking time, other moderate physical activity, and vigorous physical activity; all derived from the AAS) and quality of life. The impact of the intervention on overall quality of life (and mental health quality of life, in particular), was determined using the Assessment of Quality of Life-6D (AQoL-6D) scale [27], a 20-item instrument assessing six health-related domains. The AQoL-6D has been shown to demonstrate strong test-retest reliability (baseline and 2 weeks, $r_{\text{intraclass correlation (ICC)}}$ = .88; baseline and 1 month, r_{ICC} = .85) [28] and acceptable internal consistency (gamma coefficients [equivalent to standardized correlation coefficients] ranging from .73 to .96 for each subscale; except for sensory perception = .51) [27]. The mental health subscale has good concurrent validity when compared to the 36-item Short Form Health Survey (SF-36) (Pearson's *r*=.72) [29].

XSL•FO
RenderX

Basic demographic characteristics were also collected: date of birth, sex, highest education level (high school, post-high school [trade/certificate/diploma], university), and self-reported height and weight. The self-reported height and weight information was used to calculate body mass index (BMI), categorized as the following: underweight (<18.5 kg/m^2), normal (18.5 to <25.0 kg/m^2), overweight (25.0 to <30.0 kg/m^2), and obese (≥30.0 kg/m^2).

Participants' engagement with the app was assessed via usage statistics, including the number of visits to the app, participants' step-logging patterns, number of virtual gifts sent, and number of posts on the message walls.

Feasibility of the intervention was determined by using a purpose-designed feedback questionnaire, which was completed by intervention participants during the 8-week assessment. The scale contained nine items, each marked on a 5-point Likert scale—strongly disagree, disagree, neutral, agree, strongly agree. Three items related to perceptions of the overall app (eg, "I think the app is user friendly"), four items related to perceptions of specific features of the app (eg, "I found the daily tips useful"), and two items related to perceptions of the impact of the program (eg, "I felt like my A-Team teammates influenced me to improve my exercise regime").

Statistical Analysis

The primary outcome was change in MVPA at 8 weeks. Change in MVPA at 20 weeks and changes in all of the other outcomes at 8 and 20 weeks were considered secondary outcomes. A sample of 106 was required to detect an interaction effect size of Cohen's d=0.25 (small effect) for the primary outcome, given two groups, three repeated measures, an alpha level of .05, and 80% power (G-Power version 3.1.9.2, Universitat Kiel, Germany, 2014). The sample size was inflated to account for a design effect (potential for clustering of results within teams). Assuming an intracluster correlation coefficient of rho = .01, and approximately 5 participants per team, the design effect was 1+.01(5-1) = 1.04, therefore, the final target was 106 x 1.04 = 110 participants in total.

Participants' baseline characteristics were analyzed descriptively. Changes in primary and secondary outcomes from baseline to 8 and 20 weeks were analyzed using random-effects mixed modelling. To account for the data structure (participants nested within teams, with three repeated measures), analyses were conducted using Generalized Linear Mixed Models in SPSS version 21, with the individual and the cluster (ie, team) entered as random effects ("unstructured" covariance). The group (ie, intervention vs control), time (ie, baseline, 8 weeks, and 20 weeks), and a group x time interaction term were entered as fixed effects. The intention-to-treat principle was used for data analysis whereby all participants randomized at the commencement of the trial were retained for analysis [30]. Missing data were imputed for the small number of individuals with missing data at posttest (12/110, 10.9% of participants at 8 weeks, and 14/110, 12.7% of participants at 20 weeks) using baseline observations carried forward, which is more conservative and less susceptible to bias than last observation carried forward [31]. Where variables were right skewed (ie, physical activity variables), a log-linear distribution correction was applied.

Usage and feasibility data were analyzed descriptively using frequencies, means, and standard deviations. A small number of predefined subgroup analyses were undertaken to determine whether intervention effectiveness was related to key sociodemographic characteristics (ie, age and sex), baseline physical activity levels, and intervention dosage. Further subgroup analyses were not undertaken to prevent capitalization on chance. Baseline activity levels were categorized as sufficient or insufficient activity, according to the Australian physical activity guideline of ≥ 150 minutes of MVPA per week. Dosage was determined by dichotomizing the number of log-in occasions into low (< 18, 25/51, 49%) and high (≥ 18, 26/51, 51%) based upon a median split. The subgroup analysis was undertaken among the intervention participants only, using Generalized Linear Mixed Models, with total physical activity time entered as the target variable; individual and team entered as random effects; and age, sex, baseline adherence to MVPA guidelines, and intervention dosage entered as fixed effects. Significance for all analyses was set at $P<.05$ without adjustment for multiple comparisons, however exact P values have been reported.

Results

Participants

A total of 142 potential participants registered their interest for the study, however, only 110 were in teams that successfully formed, and were formally enrolled into the study. Of the 110 participants, 51 (46.4%) were randomized to the intervention group (12 teams) and 59 (53.6%) to the control group (13 teams). Retention at follow-up was high, with 96 of the 110 (87.3%) participants completing the 20-week follow-up. Out of 110 participants, 4 (3.6%) formally withdrew from the study citing the following reasons: needing to undergo elective surgery for a preexisting condition (intervention group, n=1), lack of time (control group, n=1), overseas vacation (control group, n=1), and too physically active (control group, n=1).

Participants' demographic and baseline characteristics are shown in Table 1. Of the 110 participants, 78 (70.9%) participants were female, and 46 (41.8%) were within the normal BMI range. There was a relatively even spread of participants across the age group categories. The majority (71/110, 64.5%) of participants were undertaking, or had completed, university education. A total of 59.1% of the sample (65/110) were insufficiently physically active (ie, achieving less than 150 min/week MVPA).

XSL•FO
RenderX

Table 1. Descriptive characteristics of the study sample at baseline (n=110).

Baseline characteristics	Intervention (n=51), n (%) or mean (SD)	Control (n=59), n (%) or mean (SD)	Total (n=110), n (%) or mean (SD)
Age in years, n (%)			
18 to <25	12 (24)	14 (24)	26 (23.6)
25 to <35	17 (33)	15 (25)	32 (29.1)
35 to <45	12 (24)	17 (29)	29 (26.4)
45 to 65	10 (20)	9 (15)	19 (17.3)
Sex, n (%)			
Male	14 (27)	12 (20)	26 (23.6)
Female	37 (73)	45 (76)	82 (74.5)
BMI [a], n (%)			
Underweight	2 (4)	1 (2)	3 (2.7)
Normal	23 (45)	23 (39)	46 (41.8)
Overweight	12 (24)	18 (31)	30 (27.3)
Obese	13 (25)	15 (25)	28 (25.5)
Highest education level, n (%)			
High school or lower	5 (10)	11 (19)	16 (14.5)
Some post-high school (eg, trade or diploma)	11 (22)	10 (17)	21 (19.1)
University	35 (69)	36 (61)	71 (64.5)
Insufficient PA[b] (<150 min/week), n (%)	33 (65)	32 (54)	65 (59.1)
Baseline total PA (min/week), mean (SD)	279 (320)	278 (313)	279 (314)
Baseline weekly total AQoL-6D[c], mean (SD)	0.80 (0.14)	0.82 (0.14)	0.81 (0.14)

[a]BMI: body mass index

[b]Physical activity (PA)—calculated as the sum of weekly walking, moderate, and vigorous physical activity.

[c]Assessment of Quality of Life-6D (AQoL-6D) scale.

Changes in Physical Activity and Quality Of Life

The results for the primary and secondary outcome measures are shown in s 2 and 3.

Both the intervention and control groups increased their MVPA time from baseline to 8 weeks (primary outcome) (see Table 2). This increase was considerably larger in magnitude for the intervention group relative to the control group—135 minutes of increase relative to the control group (treatment effect size = 0.39, *P*=.03). At 20 weeks, both groups' physical activity time remained elevated compared with baseline. Relative to the control group, the intervention group appeared to maintain a 41-minute increase (treatment effect size = 0.11), however this was not statistically significant (*P*=.26) (see Table 3).

The secondary physical activity outcomes revealed that the change in overall physical activity at 8 weeks was primarily driven by a change in time spent walking. Relative to the control group, the intervention group increased their walking time by an average of 155 minutes (treatment effect size = 0.69, *P*<.001). There were no significant group x time differences for walking at week 20, and no significant group x time effects for other types of moderate physical activity and vigorous physical activity at either time point.

There were no significant group x time effects for overall quality of life or mental health quality of life at either time point.

XSL•FO
RenderX

Table 2. Outcome measures at baseline and at 8-week follow-up.

Outcome measures	Assessment period, mean (SD)		Baseline to 8 weeks		
	Baseline	8 weeks	Mean change (SE)	Treatment effect, effect size (95% CI)	Group-by-time interaction, $F_{1,324}$ (P)
Overall PA [a] time [b] ↑ [c]					
Intervention	279 (320)	528 (391)	248 (59)	0.39 (0.01, 0.76)	4.93 (.03)
Control	278 (313)	391 (371)	113 (43)		
Walking time ↑					
Intervention	127 (198)	332 (289)	205 (38)	0.69 (0.30, 1.07)	13.01 (<.001)
Control	110 (124)	160 (185)	50 (23)		
Vigorous PA time ↑					
Intervention	52 (102)	78 (138)	26 (20)	0.12 (-0.25, 0.50)	0.89 (.35)
Control	63 (110)	83 (117)	19 (11)		
Other moderate PA time ↑					
Intervention	50 (127)	73 (154)	23 (29)	0 (-0.37, 0.38)	0.09 (.77)
Control	46 (128)	68 (171)	22 (23)		
Overall AQoL-6D [d] ↑					
Intervention	0.80 (0.14)	0.81 (0.14)	0.01 (0.01)	0.04 (-0.34, 0.41)	0.26 (.61)
Control	0.82 (0.14)	0.83 (0.14)	0.01 (0.01)		
Mental health AQoL-6D ↑					
Intervention	0.59 (0.23)	0.62 (0.23)	0.03 (0.03)	0.10 (-0.27, 0.48)	0.02 (.90)
Control	0.62 (0.25)	0.63 (0.23)	0.01 (0.02)		

[a]PA: physical activity

[b]Time is in minutes/week.

[c]Arrows (↑) indicate the desired direction (increase) of change.

[d]Assessment of Quality of Life-6D (AQoL-6D) scale.

XSL•FO
RenderX

Table 3. Outcome measures at baseline and at 20-week follow-up.

Outcome measures	Assessment period, mean (SD)		Baseline to 20 weeks		
	Baseline	20 weeks	Mean change (SE)	Treatment effect, effect size (95% CI)	Group-by-time interaction, $F_{1,324}$ (P)
Overall PA [a] time [b] ↑ [c]					
Intervention	279 (320)	376 (377)	97 (50)	0.11 (-0.26, 0.49)	1.29 (.26)
Control	278 (313)	335 (342)	56 (47)		
Walking time ↑					
Intervention	127 (198)	165 (186)	38 (29)	0.08 (-0.29, 0.46)	1.55 (.21)
Control	110 (124)	133 (137)	23 (20)		
Vigorous PA time ↑					
Intervention	52 (102)	89 (139)	37 (16)	0.15 (-0.23, 0.52)	1.41 (.24)
Control	63 (110)	82 (138)	18 (18)		
Other moderate PA time ↑					
Intervention	50 (127)	38 (100)	-12 (21)	-0.07 (-0.44, 0.31)	0.01 (.94)
Control	46 (128)	43 (95)	-3 (16)		
Overall AQoL [d] ↑					
Intervention	0.80 (0.14)	0.83 (0.15)	0.03 (0.01)	0.36 (-0.02, 0.73)	0.15 (.70)
Control	0.82 (0.14)	0.82 (0.15)	0.00 (0.01)		
Mental health AQoL ↑					
Intervention	0.59 (0.23)	0.64 (0.23)	0.05 (0.02)	0.36 (-0.02, 0.73)	0.44 (.51)
Control	0.62 (0.25)	0.61 (0.26)	-0.02 (0.02)		

[a]PA: physical activity

[b]Time is in minutes/week.

[c]Arrows (↑) indicate the desired direction (increase) of change.

[d]Assessment of Quality of Life-6D (AQoL-6D) scale.

Usage

Of the 51 participants in the intervention group, 48 (94%) used the app at least once. Usage rates were reasonably high; 28 (55%) logged steps for all 50 days of the program as intended, while 35 (69%) logged steps for 36 days or more. These steps were logged across a mean of 18 unique log-in occasions (SD 13.3, range 0-46). On average, intervention participants logged 8867 (SD 2850) steps per day, and one-third of participants (16/51, 31%) met or exceeded the intervention target of 500,000 steps in 50 days. Participants sent a mean of 4.8 gifts (SD 6.3, range 0-27) to their teammates, and made a mean of 2.7 wall posts to their team discussion wall (SD 3.4, range 0-13).

Feasibility

A total of 47 of the 51 (92%) original intervention participants completed the participant feedback questionnaire at 8 weeks. Feedback about the app overall was generally positive: 32 out of the 47 respondents (68%) either agreed or strongly agreed that the app was user friendly, 32 (68%) liked the overall presentation of the app, and 35 (74%) reported they were able to navigate easily around the app.

Feedback was also sought on the specific features of the app: 38 respondents out of 47 (81%) reported that they found the "My steps" page useful (where participants logged their daily step counts); however, there was less agreement that the daily tips were useful (18/47, 38% agreed/strongly agreed), that the virtual gifts were motivating (14/47, 30% agreed/strongly agreed), or that the unlockable awards were motivating (12/47, 26% agreed/strongly agreed).

Approximately half of the 47 respondents reported that they felt their teammates influenced them to improve their exercise regimen (27/47, 57%) and that the app provided them with social support (21/47, 45%).

Subgroup Analysis

Subgroup analyses were undertaken to determine whether, within the intervention group, change in MVPA was related to age or sex, intervention "dosage" (high vs low), and achievement of physical activity guidelines at baseline. Results showed that participants' success in the program was unrelated to sex ($F_{1,41}$= 0.10, P=.91) and age ($F_{1,41}$= 1.17, P=.32), however, it was associated with intervention dosage, with "high dose" participants increasing their MVPA significantly more than "low dose" participants ($F_{1,41}$ = 3.06, P=.04). Furthermore, participants who were insufficiently active at baseline were more likely to increase their MVPA using the program ($F_{1,41}$=

XSL•FO
RenderX

466.71, $P<.001$). Of the 33 intervention participants who were insufficiently active at baseline, 21 (64%) were sufficiently active at 8 weeks, and 13 (39%) continued to be sufficiently active at the 20-week follow-up.

No adverse events were reported throughout the trial period.

Discussion

Principal Findings

This study found that a 50-day team-based online social networking physical activity intervention incorporating pedometers produced a large and significant change in MVPA (the study's primary outcome) during the course of the intervention. The change was primarily driven by an increase in time that intervention participants spent walking (155 min/week relative to the control group). However, the intervention participants' improvements over those of the control participants were not maintained 3 months after the stimulus was removed. There was a pattern for the intervention to favorably impact on overall quality of life and mental health quality of life at the 20-week follow-up, however, this was not statistically significant. The intervention achieved reasonably high rates of engagement and retention, and participant feedback was generally positive.

To our knowledge, this is the first randomized controlled trial evaluating a physical activity intervention delivered via Facebook to report a significant improvement. Two other studies, both utilizing Facebook groups [16,32] to deliver and/or complement a physical activity intervention, reported significant time effects, but not group-by-time effects (ie, in those studies, both the intervention and control groups improved across the course of the intervention, but the degree of improvement did not significantly differ between intervention and control groups). The difference in findings may be due to the way in which Facebook was used between these two studies and our current study. Cavallo and colleagues [16] and Valle and colleagues [32] set up private Facebook groups that were intended to facilitate discussion and sharing of information between intervention participants. In contrast, our intervention involved a Facebook app (ie, standalone software, delivered via the Facebook platform) that focused on assisting participants to log, track, and compare their daily physical activity with other users. Furthermore, in both the Cavallo and colleagues [16] and Valle and colleagues [32] studies, participants who were strangers to each other offline were intended to communicate with each other via the online groups. In contrast, our study drew upon existing online social networks, so that study participants were interacting within teams of people with whom they shared an existing online connection, and presumably an offline connection as well. This is arguably more consistent with the use of online social networks, given that people typically use Facebook to interact with people with whom they share an offline connection [33]. The approach used in our study was somewhat similar to that reported by Foster and colleagues [15], whereby a group of workmates compete in a pedometer-based challenge. Similar to our study, they reported physical activity behavior change of a large magnitude.

As with many Web-based physical activity interventions [4], recidivism was apparent in the current study. At the end of the program, there was a more than 2-hour difference in weekly physical activity between groups, yet 12 weeks later at the 20-week follow-up assessments, the intervention participants' physical activity levels had returned to within 40 minutes of those of the control group (note that the study was insufficiently powered for a 40-minute difference to be statistically significant). Despite the lack of statistical significance, the trend for change provides insight for a future, larger study, as a change of this magnitude is likely to be of clinical significance. Furthermore, the subgroup analysis, which showed that around 40% of participants who were insufficiently active at baseline successfully met physical activity guidelines at the 20-week follow-up, suggests that the population impact of the app may be considerable if the intervention was to be implemented at a large scale.

Further work is required to determine how to maintain physical activity behavior change achieved by the Active Team intervention in the longer term. Fjeldsoe and colleagues' [34] review of physical activity behavior change maintenance suggests that increasing the intervention's duration, and/or building long-term follow-up prompts into the app may be useful in achieving this. While the study's intervention was mainly guided by the theory of planned behavior, other behavior change theories which emphasize behavior maintenance, such as the transtheoretical model [35] or the Health Action Process Approach [36], and self-regulation theories [37] may provide valuable insights into further strategies to maintain behavior change in the longer term.

Gamification has been a popular tech trend in recent years. The Active Team app was carefully designed to incorporate numerous gamification features; however, usage statistics and participant feedback specific to these features suggested they were not strongly embraced by participants. Despite this, the app overall achieved strong usage and participant feedback. It is possible that the influence of gamification was larger than participants indicated—that it worked in a subconscious way and did, in fact, contribute to engagement and utility of the app. Alternatively, it may be that gamification has been overhyped, or at least unsuccessful in the form in which it was implemented in our app. Such hypotheses cannot be answered by our study; indeed, the field of gamification for health behavior change is in its infancy and considerable further work is needed to explore its efficacy and optimal application.

Facebook is recognized to have extremely diverse reach, appealing to users of widely varying sociodemographic backgrounds [38,39]. Nevertheless, our study recruited a predominantly female, middle-class sample. Further work is required to determine how to attract a diverse sample, and in particular, increase reach to low physical activity/low socioeconomic status users, who are likely to gain the most benefit from a physical activity intervention. Furthermore, once effective intervention approaches have been devised, research focused on determining how to disseminate interventions on a mass scale will be key. Insights offered by social marketers and traditional marketers are likely to be highly valuable in achieving these goals.

XSL•FO
RenderX

Strengths

Strengths of the current study are the novelty of the intervention, which used online social networking to recruit participants and deliver a physical activity program (in combination with a pedometer), that the app incorporated novel features (gamification and fun), and that there was minimal contact from research personnel. The online intervention itself was delivered entirely via the software and automated emails. This hands-off delivery approach can facilitate large-scale dissemination of the intervention in the future. Further strengths of the study were the randomized controlled trial study design, and the relatively high rates of compliance and retention achieved.

Limitations

For logistical reasons, the study used self-reported measures of physical activity, and these are typically considered to be susceptible to social desirability bias [40]. Interestingly, Crutzen and Goritz [41] recently examined this issue in over 5000 participants, and found social desirability bias was, in fact, unrelated to Web-based self-reported physical activity, suggesting that Web-based self-reports of physical activity are more trustworthy and useful. An advantage of self-reported physical activity, as opposed to objectively measured physical activity, is the considerably lower participant assessment burden, which arguably enhances the study's ecological validity. Similarly, in the interest of minimizing assessment burden, we did not measure theory of planned behavior constructs and, hence, were unable to determine whether changes in these constructs explain intervention effects. In the future, a more extensive measurement protocol, including such measures, would provide useful insights into possible mechanisms. We decided not to exclude participants who stated that they obtained less than 150 minutes of weekly MVPA at enrolment, yet who went on report more than 150 minutes during the baseline surveys. In order to allow the intervention's social and team nature to function as intended, it was important to allow participants to undertake the intervention with friends, without applying too many restrictions. The application of RCT principles, such as strict eligibility criteria and prevention of contamination, in online social network interventions presents researchers with many dilemmas, and we would argue that a degree of pragmatism is required to allow the social networking intervention to function as intended, and consequently produce results that are useful in the "real world." As with most health behavior randomized controlled trials, blinding of participants to the intervention arm was not possible, however blinding of assessors was achieved since all assessments were delivered via online surveys. Additionally, this intervention had two components—the use of (1) a pedometer and (2) the app—and the use of a wait-list control meant that the individual influence of these components on study outcomes cannot be determined in the current study. Kang and colleagues' [42] meta-analysis of pedometer-based interventions found an overall effect size of 0.68 for daily step count—very similar to the effect size of 0.69 found for walking time in our study. Thus, it is possible that the pedometer component largely accounted for the behavior change observed in our study. However, engagement data indicated high use of the app, suggesting the combined intervention elements (ie, pedometer and app) were both important in achieving behavioral change. Finally, the subgroup analyses were likely underpowered, and the sample may not necessarily be generalizable given the high proportion of female and well-educated individuals.

The "snowball"-style recruitment method offers both strengths and weaknesses. Firstly, it is in keeping with how information and communication typically spread via online networks. Furthermore, it may have alleviated the problem encountered by many physical activity interventions in that they tend to attract relatively motivated individuals. It is likely that the team captains themselves may have been the "typical" motivated individuals who volunteer for research projects, however, it is plausible that the team members would not otherwise have joined, except that they received an invitation from their friend—this use of social influence is often termed "word-of-mouth seeding" in marketing [43]. An unanticipated drawback on the team structure was that numerous potential participants who registered could not be formally enrolled because their team never formed. Future iterations of the Active Team software will explore alternative recruitment structures in order to draw on the positives of snowball recruitment, without the present limitations of the strict team structure.

Conclusions

This study has provided preliminary evidence that an online, social networking physical activity intervention using pedometers can produce sizeable short-term physical activity change. Future work is needed to determine how to maintain behavior change in the longer term, how to reach underserved populations on this platform, and how to disseminate such interventions on a mass scale.

Acknowledgments

Carol Maher is funded by a fellowship from the National Heart Foundation, and this project was supported by a South Australian Premier's Research Fund Early Career Researcher Catalyst Grant. The authors thank Portal Australia for assisting with the development of the Active Team app, comedian Rachel Sommer for assisting with development of the app materials, and Emily Vaughton and Sally Turra for assisting with participant recruitment and data collection.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.2 [44].

[PDF File (Adobe PDF File), 151KB - jmir_v17i7e174_app1.pdf]

References

1. Lee I, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT, Lancet Physical Activity Series Working Group. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. Lancet 2012 Jul 21;380(9838):219-229 [FREE Full text] [doi: 10.1016/S0140-6736(12)61031-9] [Medline: 22818936]
2. Blair SN, Morris JN. Healthy hearts--and the universal benefits of being physically active: physical activity and health. Ann Epidemiol 2009 Apr;19(4):253-256. [doi: 10.1016/j.annepidem.2009.01.019] [Medline: 19344864]
3. Australian Bureau of Statistics. Canberra, Australia: Commonwealth of Australia; 2012 Oct 29. Australian Health Survey: first results, 2011-12 URL: http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4364.0.55.0012011-12?OpenDocument [accessed 2015-04-27] [WebCite Cache ID 6Y5s5G5CP]
4. Davies CA, Spence JC, Vandelanotte C, Caperchione CM, Mummery WK. Meta-analysis of Internet-delivered interventions to increase physical activity levels. Int J Behav Nutr Phys Act 2012;9:52 [FREE Full text] [doi: 10.1186/1479-5868-9-52] [Medline: 22546283]
5. Vandelanotte C, Spathonis KM, Eakin EG, Owen N. Website-delivered physical activity interventions a review of the literature. Am J Prev Med 2007 Jul;33(1):54-64. [doi: 10.1016/j.amepre.2007.02.041] [Medline: 17572313]
6. Pagoto SL, Schneider KL, Oleski J, Smith B, Bauman M. The adoption and spread of a core-strengthening exercise through an online social network. J Phys Act Health 2014 Mar;11(3):648-653. [doi: 10.1123/jpah.2012-0040] [Medline: 23416874]
7. 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]
8. Habeshian V. MarketingProfs. 2013 Apr 19. Social takes up 27% of time spent online URL: http://www.marketingprofs.com/charts/2013/10582/social-takes-up-27-of-time-spent-online [accessed 2013-09-05] [WebCite Cache ID 6JOn52rFb]
9. Schindler S. comScore. 2013 Aug 06. A digital month in Germany: what are consumers doing online? URL: http://www.comscore.com/Insights/Data-Mine/A-Digital-Month-in-Germany-What-are-Consumers-Doing-Online [accessed 2015-07-05] [WebCite Cache ID 6Zo6kucAc]
10. Facebook Newsroom. Company info URL: http://newsroom.fb.com/company-info/ [accessed 2015-04-27] [WebCite Cache ID 6Y5tnHv6M]
11. De Bruyn A, Lilien GL. A multi-stage model of word-of-mouth influence through viral marketing. International Journal of Research in Marketing 2008 Sep;25(3):151-163. [doi: 10.1016/j.ijresmar.2008.03.004]
12. Constine J. TechCrunch. 2013 Jul 24. Facebook's Q2: monthly users up 21% to 1.15B, dailies up 27% to 699M, mobile monthlies up 51% to 819M URL: http://techcrunch.com/2013/07/24/facebook-growth-2/ [accessed 2015-07-02] [WebCite Cache ID 6ZiXx9nqe]
13. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining gamification. In: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments. New York, NY: ACM; 2011 Presented at: 15th International Academic MindTrek Conference: Envisioning Future Media Environments; September 28-30, 2011; Tampere, Finland p. 9-15.
14. Takahashi D. VentureBeat. 2010 Aug 25. Website builder DevHub gets users hooked by "gamifying" its service URL: http://venturebeat.com/2010/08/25/devhub-scores-engagement-increase-by-gamifying-its-web-site-creation-tools/ [accessed 2015-04-27] [WebCite Cache ID 6Y5uDwnQQ]
15. Foster D, Linehan C, Kirman B, Lawson S, James G. Motivating physical activity at work: using persuasive social media for competitive step counting. In: Proceedings of the 14th International Academic MindTrek Conference: Envisioning Future Media Environments. New York, NY: ACM; 2010 Presented at: 14th International Academic MindTrek Conference: Envisioning Future Media Environments; October 6-8, 2010; Tampere, Finland p. 111-116.
16. Cavallo DN, Tate DF, Ries AV, Brown JD, DeVellis RF, Ammerman AS. A social media-based physical activity intervention: a randomized controlled trial. Am J Prev Med 2012 Nov;43(5):527-532 [FREE Full text] [doi: 10.1016/j.amepre.2012.07.019] [Medline: 23079176]
17. Napolitano MA, Hayes S, Bennett GG, Ives AK, Foster GD. Using Facebook and text messaging to deliver a weight loss program to college students. Obesity (Silver Spring) 2013 Jan;21(1):25-31. [doi: 10.1002/oby.20232] [Medline: 23505165]
18. Turner-McGrievy G, Tate D. Tweets, apps, and pods: results of the 6-month Mobile Pounds Off Digitally (Mobile POD) randomized weight-loss intervention among adults. J Med Internet Res 2011;13(4):e120 [FREE Full text] [doi: 10.2196/jmir.1841] [Medline: 22186428]
19. Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, CONSORT GROUP (Consolidated Standards of Reporting Trials). The revised CONSORT statement for reporting randomized trials: explanation and elaboration. Ann Intern Med 2001 Apr 17;134(8):663-694. [Medline: 11304107]

XSL•FO
RenderX

20. Kernot J, Olds T, Lewis LK, Maher C. Usability testing and piloting of the Mums Step It Up program--a team-based social networking physical activity intervention for women with young children. PLoS One 2014;9(10):e108842 [FREE Full text] [doi: 10.1371/journal.pone.0108842] [Medline: 25272039]
21. Tudor-Locke C, Craig CL, Aoyagi Y, Bell RC, Croteau KA, De Bourdeaudhuij I, et al. How many steps/day are enough? For older adults and special populations. Int J Behav Nutr Phys Act 2011;8:80 [FREE Full text] [doi: 10.1186/1479-5868-8-80] [Medline: 21798044]
22. Ajzen I. From intentions to actions: a theory of planned behavior. In: Kuhl J, Beckmann J, editors. Action Control: From Cognition to Behavior. Heidelberg, Germany: Springer Berlin Heidelberg; 1985:11-39.
23. Ajzen I. Attitudes, traits and actions: dispositional prediction of behaviour in personality and social psychology. In: Berkowitz L, editor. Advances in Experimental Social Psychology. Volume 20. San Diego, CA: Academic Press, Inc; 1987:1-63.
24. Dodd KJ, Taylor NF, Graham HK. A randomized clinical trial of strength training in young people with cerebral palsy. Dev Med Child Neurol 2003 Oct;45(10):652-657. [Medline: 14515935]
25. AIHW. Australian Institute of Health and Welfare (AIHW). 2003 Apr 24. The Active Australia Survey: a guide and manual for implementation, analysis and reporting URL: http://www.aihw.gov.au/publication-detail/?id=6442467449 [accessed 2015-04-27] [WebCite Cache ID 6Y5vCGjpE]
26. Brown WJ, Burton NW, Marshall AL, Miller YD. Reliability and validity of a modified self-administered version of the Active Australia physical activity survey in a sample of mid-age women. Aust N Z J Public Health 2008 Dec;32(6):535-541. [doi: 10.1111/j.1753-6405.2008.00305.x] [Medline: 19076744]
27. Richardson JR, Peacock SJ, Hawthorne G, Iezzi A, Elsworth G, Day NA. Construction of the descriptive system for the Assessment of Quality of Life AQoL-6D utility instrument. Health Qual Life Outcomes 2012;10:38 [FREE Full text] [doi: 10.1186/1477-7525-10-38] [Medline: 22507254]
28. Richardson J, Chen G, Iezzi A, Khan MA. Transformations Between the Assessment of Quality of Life AQoL Instruments and Test-Retest Reliability. Melbourne, Australia: Centre for Health Economics, Monash University; 2011 Jun. URL: http://www.buseco.monash.edu.au/centres/che/pubs/researchpaper66.pdf [accessed 2015-04-27] [WebCite Cache ID 6Y5vVW2rz]
29. Allen J, Inder KJ, Lewin TJ, Attia JR, Kelly BJ. Construct validity of the Assessment of Quality of Life - 6D (AQoL-6D) in community samples. Health Qual Life Outcomes 2013;11:61 [FREE Full text] [doi: 10.1186/1477-7525-11-61] [Medline: 23590808]
30. Campbell MK, Piaggio G, Elbourne DR, Altman DG. Consort 2010 statement: extension to cluster randomised trials. BMJ 2012;345:e5661. [Medline: 22951546]
31. Obesity and the Facts: An Analysis of Data From the Health Survey for England 2003. Oxford, UK: Social Issues Research Centre; 2005 Feb. URL: http://www.sirc.org/obesity/obesityfacts.pdf [accessed 2015-04-27] [WebCite Cache ID 6Y5w6v6no]
32. Valle CG, Tate DF, Mayer DK, Allicock M, Cai J. A randomized trial of a Facebook-based physical activity intervention for young adult cancer survivors. J Cancer Surviv 2013 Sep;7(3):355-368 [FREE Full text] [doi: 10.1007/s11764-013-0279-5] [Medline: 23532799]
33. Ellison N, Steinfield C, Lampe C. The benefits of Facebook "friends:" social capital and college students' use of online social network sites. J Comput Mediat Commun 2007 Jul;12(4):1143-1168. [doi: 10.1111/j.1083-6101.2007.00367.x]
34. Fjeldsoe B, Neuhaus M, Winkler E, Eakin E. Systematic review of maintenance of behavior change following physical activity and dietary interventions. Health Psychol 2011 Jan;30(1):99-109. [doi: 10.1037/a0021974] [Medline: 21299298]
35. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot 1997;12(1):38-48. [Medline: 10170434]
36. Schwarzer R, Luszczynska A. How to overcome health-compromising behaviors. Eur Psychol 2008 Jan 3;13(2):141-151. [doi: 10.1027/1016-9040.13.2.141]
37. Maes S, Karoly P. Self-regulation assessment and intervention in physical health and illness: A review. Appl Psychol 2005 Apr 19;54(2):267-299 [FREE Full text] [doi: 10.1111/j.1464-0597.2005.00210.x]
38. Duggan M, Brenner J. Pew Research Center. Washington, DC: Pew Internet & American Life Project; 2013 Feb 14. The demographics of social media users — 2012 URL: http://www.pewinternet.org/2013/02/14/the-demographics-of-social-media-users-2012/ [accessed 2015-04-27] [WebCite Cache ID 6Y5wqtPXs]
39. Guadagno R, Muscanell N, Pollio D. The homeless use Facebook?! Similarities of social network use between college students and homeless young adults. Comput Human Behav 2013;29(1):86-89. [doi: 10.1016/j.chb.2012.07.019]
40. Motl RW, McAuley E, DiStefano C. Is social desirability associated with self-reported physical activity? Prev Med 2005 Jun;40(6):735-739. [doi: 10.1016/j.ypmed.2004.09.016] [Medline: 15850873]
41. Crutzen R, Göritz AS. Does social desirability compromise self-reports of physical activity in Web-based research? Int J Behav Nutr Phys Act 2011;8:31 [FREE Full text] [doi: 10.1186/1479-5868-8-31] [Medline: 21492435]
42. Kang M, Marshall SJ, Barreira TV, Lee J. Effect of pedometer-based physical activity interventions: a meta-analysis. Res Q Exerc Sport 2009 Sep;80(3):648-655. [doi: 10.1080/02701367.2009.10599604] [Medline: 19791652]
43. Libai B, Muller E, Peres R. Decomposing the value of word-of-mouth seeding programs: acceleration versus expansion. J Mark Res 2013 Apr;50(2):161-176.

XSL•FO
RenderX

44. 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

AAS: Active Australia Survey
AQoL-6D: Assessment of Quality of Life-6D
BMI: body mass index
ICC: intraclass correlation
MVPA: moderate-to-vigorous physical activity
PA: physical activity
RCT: randomized controlled trial
SF-36: 36-item Short Form Health Survey

Edited by G Eysenbach; submitted 30.11.14; peer-reviewed by R Maddison, D Cavallo, E Hekler; comments to author 12.02.15; revised version received 27.04.15; accepted 24.05.15; published 13.07.15.

Please cite as:
Maher C, Ferguson M, Vandelanotte C, Plotnikoff R, De Bourdeaudhuij I, Thomas S, Nelson-Field K, Olds T
A Web-Based, Social Networking Physical Activity Intervention for Insufficiently Active Adults Delivered via Facebook App: Randomized Controlled Trial
J Med Internet Res 2015;17(7):e174
URL: http://www.jmir.org/2015/7/e174/
doi:10.2196/jmir.4086
PMID:26169067

©Carol Maher, Monika Ferguson, Corneel Vandelanotte, Ron Plotnikoff, Ilse De Bourdeaudhuij, Samantha Thomas, Karen Nelson-Field, Tim Olds. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 13.07.2015. 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

Engagement and Nonusage Attrition With a Free Physical Activity Promotion Program: The Case of 10,000 Steps Australia

Diana Guertler[1,2,3], Dipl-Psych; Corneel Vandelanotte[2], PhD; Morwenna Kirwan[4], PhD; Mitch J Duncan[2,5], PhD

[1]Institute of Social Medicine and Prevention, University Medicine, Greifswald, Germany

[2]School of Human Health and Social Sciences, Physical Activity Research Group, Rockhampton, Australia

[3]DZHK (German Centre for Cardiovascular Research), Greifswald partner site, Greifswald, Germany

[4]School of Science and Health, University of Western Sydney, Sydney, Australia

[5]School of Medicine & Public Health, Priority Research Centre for Physical Activity and Nutrition, Faculty of Health and Medicine, The University of Newcastle, Newcastle, Australia

Corresponding Author:
Diana Guertler, Dipl-Psych
Institute of Social Medicine and Prevention
University Medicine
Walther-Rathenau-Strasse 48
Greifswald, 17475
Germany
Phone: 49 3834867716
Fax: 49 3834867701
Email: diana.guertler@uni-greifswald.de

Abstract

Background: Data from controlled trials indicate that Web-based interventions generally suffer from low engagement and high attrition. This is important because the level of exposure to intervention content is linked to intervention effectiveness. However, data from real-life Web-based behavior change interventions are scarce, especially when looking at physical activity promotion.

Objective: The aims of this study were to (1) examine the engagement with the freely available physical activity promotion program 10,000 Steps, (2) examine how the use of a smartphone app may be helpful in increasing engagement with the intervention and in decreasing nonusage attrition, and (3) identify sociodemographic- and engagement-related determinants of nonusage attrition.

Methods: Users (N=16,948) were grouped based on which platform (website, app) they logged their physical activity: Web only, app only, or Web and app. Groups were compared on sociodemographics and engagement parameters (duration of usage, number of individual and workplace challenges started, and number of physical activity log days) using ANOVA and chi-square tests. For a subsample of users that had been members for at least 3 months (n=11,651), Kaplan-Meier survival curves were estimated to plot attrition over the first 3 months after registration. A Cox regression model was used to determine predictors of nonusage attrition.

Results: In the overall sample, user groups differed significantly in all sociodemographics and engagement parameters. Engagement with the program was highest for Web-and-app users. In the subsample, 50.00% (5826/11,651) of users stopped logging physical activity through the program after 30 days. Cox regression showed that user group predicted nonusage attrition: Web-and-app users (hazard ratio=0.86, 95% CI 0.81-0.93, P<.001) and app-only users (hazard ratio=0.63, 95% CI 0.58-0.68, P<.001) showed a reduced attrition risk compared to Web-only users. Further, having a higher number of individual challenges (hazard ratio=0.62, 95% CI 0.59-0.66, P<.001), workplace challenges (hazard ratio=0.94, 95% CI 0.90-0.97, P<.001), physical activity logging days (hazard ratio=0.921, 95% CI 0.919-0.922, P<.001), and steps logged per day (hazard ratio=0.99999, 95% CI 0.99998-0.99999, P<.001) were associated with reduced nonusage attrition risk as well as older age (hazard ratio=0.992, 95% CI 0.991-0.994, P<.001), being male (hazard ratio=0.85, 95% CI 0.82-0.89, P<.001), and being non-Australian (hazard ratio=0.87, 95% CI 0.82-0.91, P<.001).

Conclusions: Compared to other freely accessible Web-based health behavior interventions, the 10,000 Steps program showed high engagement. The use of an app alone or in addition to the website can enhance program engagement and reduce risk of attrition. Better understanding of participant reasons for reducing engagement can assist in clarifying how to best address this issue to maximize behavior change.

XSL•FO
RenderX

(J Med Internet Res 2015;17(7):e176) doi:10.2196/jmir.4339

KEYWORDS

physical activity; Internet; engagement; smartphone

Introduction

Insufficient physical activity has been identified as the fourth-leading risk factor for global mortality [1]. Previous research has shown that Web-based behavior change interventions can be effective in increasing physical activity [2,3]. The Internet is a promising tool to deliver complex, individualized, and tailored interventions while reaching a large part of the population at lower cost than face-to-face interventions [4]. However, maintaining interest of the participants in the intervention over time has been reported as a main challenge of Internet-delivered interventions [5]. Web-based intervention studies typically suffer from high nonusage attrition (ie, not all participants use or keep using the intervention as intended by the developers) [6]. Most often, website log-ins, a frequently used indicator of engagement and intervention exposure, decrease rapidly over time [7,8]. This makes it difficult to measure intervention effects because participants receive different doses of the intervention content [5]. For example, in a Web-based weight loss study [9], only 64% of the intervention group actually used the intervention at least once. This is important because the level of exposure to intervention content has been linked to intervention effectiveness [10,11]. In addition to nonusage attrition, Web-based interventions studies suffer also from high dropout attrition (ie, participants are lost to follow-up). For Web-based physical activity interventions, reported dropout attrition rates vary between 0% and 62% [2,12].

Eysenbach [5] calls for a science of attrition to systematically examine attrition rates, engagement measures, and associated variables. Thus far, research on attrition has been done primarily in the context of controlled trials [6,13]. However, to evaluate the real public health impact of an intervention there is a need to examine effectiveness and reach of the target population after implementation to real-life settings [14]. Findings of nonusage rates from controlled settings may not translate to real-life settings. In efficacy trials, participants usually have gone through a rigorous screening process to determine eligibility and therefore include a selected group of participants that is likely more motivated to use the intervention compared to those not undergoing this screening. Further, people could be more committed to the study because of the formal structure of the trial or active recruitment, which leads to lower attrition [5,6]. This is supported by studies that find a higher percentage of intervention completers and higher website usage for trial users compared to "real-life" users of the same website [15,16]. Although dropout and nonusage attrition have been described in relation to commercially available websites [17,18], few studies describe similar patterns in freely accessible interventions, especially in regard to physical activity [15].

Several intervention characteristics have shown to enhance engagement and/or decrease attrition of interventions including the provision of personally tailored content, interactive components, social networking, and reminders [19-22]. Besides characteristics of the intervention itself, personal characteristics of the users and the degree of engagement with the intervention may affect nonusage attrition [15,20,23]. Attempts to promote physical activity via smartphone technology appear to be promising because it increases the convenience of accessing the intervention and engaging in self-monitoring [24]. Within controlled trials, there is some evidence that using smartphone apps can enhance engagement, decrease attrition, and increase efficacy of Web-based interventions [25-27]. However, thus far there is no knowledge about how smartphone apps can enhance engagement with Web-based interventions in real-life settings.

The aim of this study is to examine engagement with a Web-based physical activity intervention in real life because findings from controlled settings may not translate into real-life settings. Therefore, this study examines engagement and nonusage attrition in the freely available 10,000 Steps Australia program [20,28], which aims to promote physical activity through the use of pedometers and a website. A second aim is to examine whether use of a smartphone app is associated with reduced nonusage attrition and increased engagement with the intervention. Third, we aim to identify sociodemographic- and engagement-related determinants of nonusage attrition.

Methods

Intervention Program

The 10,000 Steps program is a freely available physical activity promotion program that encourages users to record and monitor their physical activity using pedometers. It was initially developed as a whole-community multilevel program based on the socioecological framework. Further information on the development of the program has been reported elsewhere [20,28]. A main feature of the program is an online step log that allows users to enter and monitor their daily physical activity levels based on pedometer steps or time spent in physical activity. This feature is available to users both on the website and as a smartphone app. Further, users are able to join individual challenges where they choose from a monthly updated selection of goals and receive graph- and text-based feedback of their progress (individual challenge). When users are recruited via their workplace, they may participate in team-based workplace challenges. These usually last longer than 1 month and the workplace is responsible to set the team challenges (workplace challenge). Further, users are able to use a discussion forum and virtual walking buddies with whom they can share progress.

Smartphone App

In addition to the opportunity to use the program via the 10,000 Steps website, there is also a smartphone app available on the iOS mobile platform [27]. Initially designed to allow users to enter their daily physical activity, users are now also able to join and view progress of challenges. Data from the smartphone

XSL•FO
RenderX

app are synchronized with their activities recorded on the website. Data from a case-matched control trial [27] indicated that use of the 10,000 Steps app increased the number of days physical activity was logged and the likelihood to log more than 10,000 steps per day in a sample of 10,000 Steps users. Excluding participants from the controlled trial, the app had a total of 35,761 downloads as of April 30, 2014. One-third (33.12%, 11,845/35,761) of those that downloaded the app also used it to log physical activity.

Data Collection and Extraction

Data were extracted from the database of the 10,000 Steps program for users aged at least 18 years, who registered between July 8, 2013 and April 30, 2014, and who logged physical activity for at least 1 day in this period either through the website or the smartphone app (N=17,590). Although the website and app were available before July 8, 2013, this date was chosen as the start date for data extraction because prior to this date the available version of the app had reduced functionality; specifically, it did not allow users to join or view progress of challenges. No major changes in app functionality have occurred since this date, only minor updates have happened. Therefore, this study was delimited to users who registered during this 297-day period between July 8, 2013 and April 30, 2014 excluding those users who registered before this time. Thus, the maximum membership length in this study was 297 days. Website and app usage data were automatically recorded while using the 10,000 Steps program. When registering with the 10,000 Steps program, participants provided informed consent for usage of their data for research purposes.

Measures

Sociodemographics, "Where Did You Hear About Us?," and Length of Membership

Date of birth, gender, and country of residence were assessed when participants registered to the 10,000 Steps program. Users were also asked how they heard about the 10,000 Steps program (21 response options were provided, including different types of media, friends, workplace, health professionals, and specific initiatives the program was advertised in). Length of membership was calculated as the number of days between the date of registration for the program and April 30, 2014.

Engagement Parameters

Engagement was defined as the duration and frequency of involvement with the program. Four measures of engagement were used: (1) the duration of program use calculated as the number of calendar days from the first to the last time the physical activity log was used, (2) the number of individual challenges initiated, (3) the number of workplace challenges initiated, and (4) the total number of days physical activity was recorded in the step log (both website and app). The number of individual and workplace challenges participated in was determined from the 10,000 Steps database, which encompassed website and app usage information.

The total number of days of physical activity recorded in the step log and duration of program usage differed. For example, a user may have used the program for 50 days (time from first to last step log), but only logged steps on 30 occasions during this time.

Nonusage Attrition

Duration of program usage was also applied as an indicator of nonusage attrition. Participants were coded as "nonusage attrition was observed" when they did not log physical activity for at least 14 days [29] (ie, there were ≥14 days between their last physical activity log and the end of the observation period). All other users were coded as "nonusage attrition was not observed." For example, when a participant first logged steps on the 10,000 Steps platform on October 1, 2013 and the last time on December 1, 2013, then nonusage attrition was deemed to have occurred after 62 days of use.

Physical Activity

Users' mean steps per day were determined by dividing the total number of steps logged by the number of days steps were logged for. If users' mean of logged steps per occasion was more than 20,000, this was truncated to 20,000 steps [30,31]. Truncation was performed for 928 participants.

Data Analysis

Overview

The program allows participants to retrospectively log steps in case they started to use a pedometer before registering with the program. Users were excluded from data analysis when they logged physical activity for more than 7 days prior to their start of their membership (n=617) and when inconsistencies in data were detected; that is, users logged on average less than 100 steps per day (n=16) or logged physical activity on more days than it would be possible based on their duration of program use (n=9). The final number of users included in the analysis of user characteristics and engagement was N=16,948 and is referred to as the overall sample. Attrition analysis was based on a subsample of users as described subsequently. Based on the platform used to log steps, 3 groups were defined: Web-only users who logged steps solely via the 10,000 Steps website (83.87%, 14,215/16,948), app-only users who logged steps solely via the 10,000 Steps smartphone app (8.56%, 1451/16,948), and Web-and-app users who logged steps via the 10,000 Steps website and the smartphone app (7.56%, 1282/16,948).

User Characteristics and Intervention Engagement

First, means and standard deviations for sociodemographics, engagement parameters, and logged physical activity were calculated to describe the overall sample and the 3 user groups.

To assess which personal characteristics of 10,000 Steps users may facilitate choosing the app over the website, differences between the user groups regarding sociodemographics were analyzed. Further, user groups were compared regarding engagement parameters and logged physical activity. Group comparisons were performed using 1-way ANOVAs with Bonferroni-corrected post hoc comparisons (for continuous variables) and chi-square tests (for categorical variables).

Because engagement parameters are likely to depend on sociodemographics, length of membership, and physical activity

XSL•FO
RenderX

level, we also examined the effect of user group on the engagement parameters using linear regression adjusting for those variables. Four regression analyses were fitted, each using an engagement parameter (duration of usage, individual challenges, workplace challenges, physical activity log days) as the dependent variable. User group was used as the independent variable along with age, gender, country of residency, length of membership, and physical activity as covariates. Standardized regression coefficients were calculated. Effect sizes (η^2) were calculated for linear regression because even small differences tend to reach significance with high numbers of participants such as in our study. According to Cohen [32], the minimum criterion for at least a small effect was η^2>.01. Therefore, in this study, differences were considered meaningful when effect sizes reached these thresholds. Further, duration of usage, number of individual and workplace challenges started, and number of days physical activity was logged for were presented for different membership lengths.

Nonusage Attrition

Survival analysis was used to examine differences in nonusage attrition between groups. Nonusage attrition was examined over the first 3 months after registration to the program and was limited to a subsample of users (n=11,651) who were a 10,000 Steps member for at least 3 months (90 days) at the time of data extraction (April 30, 2014). This was done to ensure that all participants had the same chance to use the program and to enable comparability with other published attrition curves [5,15,33]. This means that users were included even if they only logged steps for a single day, but had been a member for 3 months. The duration of program use was used as the time variable. The event variable was coded as specified in the Measures section with 1=nonusage attrition observed and 0=nonusage attrition not observed. Kaplan-Meier survival curves showing the proportion of users surviving over time and quartiles of survival time were estimated by user group. The equality of the survivor functions was tested with a log-rank test. Predictors of nonusage attrition (user group, sociodemographics, engagement parameters, and steps per day) were examined within univariate Cox proportional hazard regression. Predictors that had a univariate *P* value of <.25 [34] were selected for inclusion in a multivariate model. Hazard ratios, which represent relative risks for attrition, were calculated. Statistical analysis was performed with Stata version 12 (StataCorp LP, College Station, TX, USA).

Results

Overview

Of 16,040 users who answered the “Where did you hear about us?” question, most users (73.99%, 11,868/16,040) indicated that they heard about the program through their workplace. Further, 12.42% (1992/16,040) heard about 10,000 Steps from a friend; 4.3% (692/16,040) from a webpage; 2.9% (461/16,040) from a health professional; 0.9% (140/16,040) from Facebook; 0.41% (65/16,040) from other media including TV, newspaper, and radio; and 5.1% (822/16,040) indicated other sources.

User Characteristics

Descriptions of the overall sample and the subsample regarding sociodemographics, engagement data, and logged physical activity are shown in Table 1 and Multimedia Appendix 1, respectively. The majority of participants in the overall sample were female (69.87%, 11,841/16,948) and Australian (77.54%, 13,142/16,948). Membership length ranged between 1 and 297 days (mean 190, SD 78.6 days). Gender and country of residence were significantly different between user groups, with highest percentage of females (73.5%, 942/1282) and Australians (84.7%, 1086/1282) for the Web-and-app group. App-only users and Web-and-app users were significantly younger than Web-only users (*P*<.001) with app-only users being also younger than Web-and-app users (*P*<.04). Web-only users (*P*=.004) and Web-and-app users (*P*<.001) had longer durations of membership in the 10,000 Steps program compared to the app-only users, with no difference between Web-only and Web-and-app users (*P*=.28).

XSL•FO
RenderX

Table 1. Sociodemographics, engagement, and physical activity by user group in the overall sample (N=16,948).

Variables	Overall N=16,948	Web only n=14,215	App only n=1451	Web and app n=1282	$F_{2,2}$	χ^2_2	*P*
Sociodemographics							
Age, mean (SD)	41.8 (12.1)	42.4 (12.2)	38.3 (11.1)	39.5 (11.5)	102.6		<.001[abc]
Females, n (%)	11,841 (69.87)	9848 (69.28)	1051 (72.43)	942 (73.48)		14.8	<.001
Australians, n (%)	13,142 (77.54)	10,936 (76.93)	1120 (77.19)	1086 (84.71)		41.0	<.001
Membership days, mean (SD)	190.4 (78.6)	190.7 (78.9)	183.9 (79.6)	194.6 (72.7)	7.0		<.001[ac]
Engagement							
Duration of usage (days), mean (SD)	34.5 (30.5)	32.8 (28.2)	37.7 (37.2)	50.2 (40.4)	203.6		<.001[abc]
Individual challenges, mean (SD)	0.1 (0.5)	0.1 (0.4)	0.2 (0.7)	0.3 (0.9)	174.4		<.001[abc]
Workplace challenges, mean (SD)	0.9 (0.5)	0.9 (0.5)	0.8 (0.7)	0.9 (0.6)	52.2		<.001[ac]
Number of days physical activity was logged for, mean (SD)	30.8 (25.1)	29.5 (23.6)	32.9 (29.8)	43.3 (31.5)	188.7		<.001[abc]
Physical activity							
Steps per day, mean (SD)	10,692.1 (4194.4)	10,701.7 (4251.6)	10,253.0 (3951.1)	11,082.6 (3758.1)	13.6		<.001[abc]

[a] Web only is different from app only.

[b] Web only is different from Web and app.

[c] App only is different from Web and app.

Engagement With the Intervention

In the overall sample, users utilized the program between 1 and 296 days (mean 34.5, SD 30.5 days) with 6.51% (1103/16,948) and 81.89% (13,879/16,948) participating at least in 1 individual challenge (range 0-8) and 1 workplace challenge (range 0-8), respectively. Users logged on average 30.8 days of physical activity (range 1-290 days) with 97.77% (16,400/16,948) of users logging physical activity more than once. With increasing length of membership, the average duration of usage, the number of individual and workplace challenges, as well as the number of days physical activity was logged per week decreased (Table 2). For example, users who were members between 1 and 2 months (30-60 days) used the program a mean 4.1 (SD 2.4) days per week, whereas users with a membership of at least 9 months (278-297 days) used the program a mean 1.1 (SD 0.9) days per week.

All engagement parameters differed significantly across user groups (*P*<.001). App-only users showed a longer duration of usage (*P*<.001), a higher number of individual challenges (*P*<.001), a lower number of workplace challenges (*P*<.001), and a higher number of days they logged physical activity (*P*<.001) compared to Web-only users (Table 1). Compared to app-only and Web-only users, Web-and-app users had a longer duration of usage (*P*<.001), higher number of individual (*P*=.02 and *P*<.001, respectively) challenges, as well as a higher number of days they logged physical activity (*P*<.001). Regarding workplace challenges, Web-and-app users had higher numbers compared to app-only users (*P*<.001), but were not significantly different from Web-only users (*P*=.07).

XSL•FO
RenderX

Table 2. Mean engagement parameters for different membership lengths in the overall sample (N=16,948).

Membership length	n	Duration of platform usage (days)		Individual challenges		Workplace challenges		Days physical activity logged	
		Mean (SD)	Mean/week[a](SD)	Mean (SD)	Mean/week[a](SD)	Mean (SD)	Mean/week[a](SD)	Mean (SD)	Mean/week[a](SD)
≤1 week (1-7 days)	54	3.6 (2.7)	10.1[b](10.7)	0.00 (0.00)	0.00 (0.00)	0.24 (0.43)	0.59 (1.46)	3.5 (2.6)	9.9[b] (10.6)
1-2 weeks (8-14 days)	60	9.0 (5.5)	5.5 (3.4)	0.10 (0.35)	0.07 (0.24)	0.60 (0.49)	0.37 (0.33)	8.7 (5.3)	5.4 (3.2)
2-3 weeks (15-21 days)	234	13.0 (6.1)	5.1 (2.4)	0.09 (0.09)	0.00 (0.04)	0.87 (0.46)	0.34 (0.18)	12.5 (5.9)	4.9 (2.3)
3-4 weeks (22-29 days)	413	15.7 (7.7)	4.6 (2.3)	0.04 (0.22)	0.01 (0.06)	0.85 (0.48)	0.25 (0.15)	14.0 (8.0)	4.1 (2.3)
1-2 months (30-60 days)	874	23.8 (14.4)	4.1 (2.4)	0.09 (0.35)	0.01 (0.06)	0.63 (0.49)	0.11 (0.09)	22.5 (14.1)	3.8 (2.4)
2-3 months (61-91 days)	1592	34.1 (18.7)	3.2 (1.8)	0.11 (0.41)	0.01 (0.04)	0.89 (0.50)	0.08 (0.04)	31.7 (17.8)	3.0 (1.7)
3-4 months (92-122 days)	604	36.2 (28.9)	2.4 (1.9)	0.26 (0.73)	0.02 (0.05)	0.66 (0.59)	0.04 (0.04)	32.8 (26.9)	2.2 (1.8)
4-5 months (123-153 days)	370	38.7 (32.5)	1.9 (1.6)	0.19 (0.67)	0.01 (0.03)	0.71 (0.65)	0.04 (0.03)	33.5 (28.6)	1.7 (1.4)
5-6 months (154-184 days)	1335	35.0 (34.0)	1.4 (1.4)	0.12 (0.56)	0.01 (0.02)	0.87 (0.53)	0.03 (0.02)	29.2 (24.5)	1.2 (1.0)
6-7 months (185-215 days)	3218	31.6 (27.9)	1.1 (1.0)	0.10 (0.47)	0.00 (0.02)	0.92 (0.37)	0.03 (0.01)	28.2 (22.7)	1.0 (0.8)
7-8 months (216-246 days)	4016	36.7 (34.1)	1.1 (1.0)	0.12 (0.61)	0.00 (0.02)	0.93 (0.65)	0.03 (0.02)	32.7 (27.5)	1.0 (0.8)
8-9 months (247-277 days)	2462	36.8 (30.5)	1.0 (0.8)	0.09 (0.49)	0.00 (0.01)	0.91 (0.46)	0.02 (0.01)	33.4 (26.0)	0.9 (0.7)
9-10 months (278-297 days)	1716	45.1 (37.4)	1.1 (0.9)	0.08 (0.42)	0.00 (0.01)	1.0 (0.58)	0.02 (0.01)	38.9 (30.4)	0.9 (0.7)

[a] Calculated by (number of days respectively challenges / individual membership days)*7.

[b] Numbers exceed 7 days (maximum membership length in this group) due to the opportunity to retrospectively log steps before registration.

Recorded Physical Activity

More than half of participants (53.30%, 9033/16,948) logged at least 10,000 steps on average per day. However, app-only users had significantly lower numbers of steps per day compared to Web-only and Web-and-app users (P<.001) with lower numbers of steps per day for Web-only users compared to Web-and-app users (P=.01). Web-and-app users logged a mean of 67.4% (SD 29.94) of their total steps through the app.

Prediction of Engagement Parameters by Group

Table 3 shows results of 4 linear regression analyses regarding the prediction of engagement parameters by user group when controlling for sociodemographics, length of membership, and logged physical activity. Results align with data from Table 1: comparisons of the number of workplace challenges within user groups did not reach the threshold for a meaningful effect when using Web-only as reference category (B=–0.07, η^2=.005 and B=0.01, η^2=.000). Web-and-app users showed a longer duration of usage (B=0.16, η^2=.026), more individual challenges (B=0.12, η =.014), and more days of physical activity logged (B=0.15, η^2=.024) compared to Web-only users; however, comparisons between app-only and Web-only users regarding duration of usage, number of individual challenges, and physical activity log days did not reach the threshold for a meaningful effect (η^2=.004, η^2=.009, η^2=.003, respectively).

XSL•FO
RenderX

Table 3. Linear regression analyses showing associations between engagement parameters and app-only and Web-and-app groups in comparison to Web only in the overall sample (N=16,948).

Dependent variables	App only[a]			Web and app[a]		
	B (SE)	*P*	η^2	B (SE)	*P*	η^2
Duration of usage	0.06 (0.82)	<.001	.004	0.16 (0.86)	<.001	.026
Individual challenges	0.10 (0.01)	<.001	.009	0.12 (0.02)	<.001	.014
Workplace challenges	–0.07 (0.02)	<.001	.005	0.01 (0.02)	.14	.000
Physical activity log days	0.06 (0.67)	<.001	.003	0.15 (0.71)	<.001	.024

[a] Web only was used as reference category; analyses controlled for age, gender, country, length of membership, and physical activity logged.

Nonusage Attrition

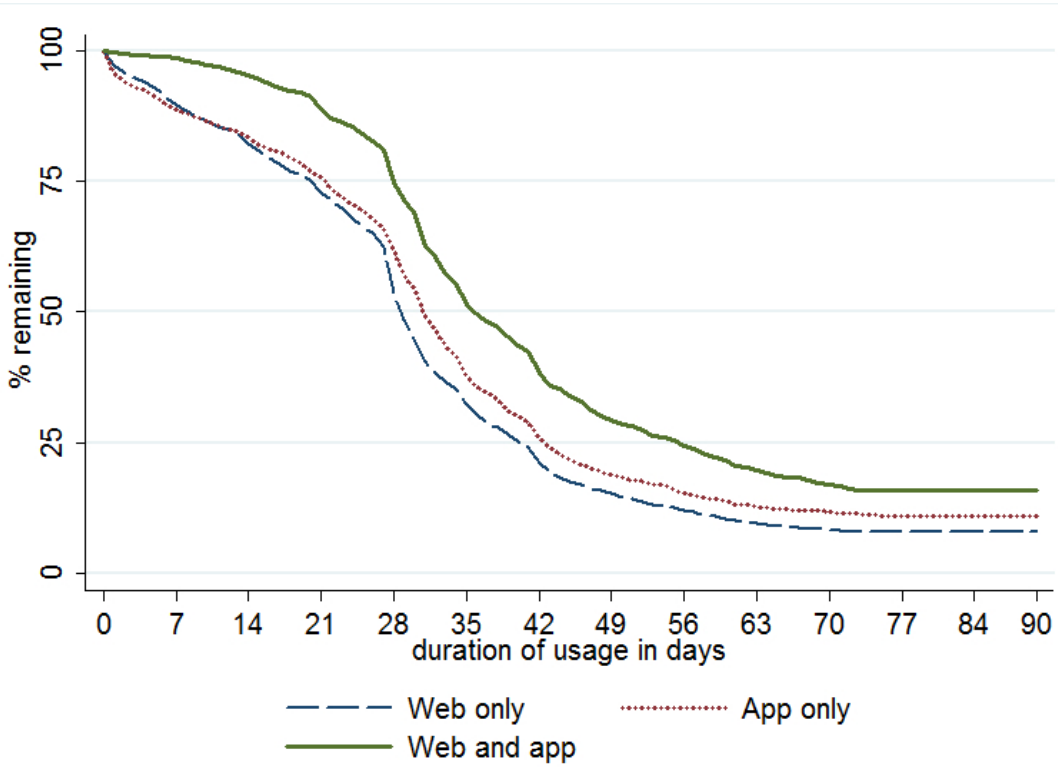
The following results are based on a subsample only including users that had been a 10,000 Steps member for at least 3 months (n=11,651). Figure 1 presents Kaplan-Meier survival curves for the different user groups based on the duration of usage. The log-rank test showed that the survivor functions were significantly different across groups (χ^2_2=161.3, *P*<.001). Estimated median lifetime usage (time after which 50% stopped logging physical activity) was 30 days for all groups combined (Table 4). For all groups combined, 25.00% (2913/11,651) were still logging steps after 42 days. This was similar to the Web-only and app-only groups, with 41 and 43 days. respectively; however, in the Web-and-app group, 25.0% (220/878) of the sample were still logging steps after 56 days.

Table 4. Survival time by group in the subsample of users who were 10,000 Steps members for at least 3 months (n=11,651).

User group	Percentage of group still used platform[a]		
	75%	50%	25%
Web only	21 days	29 days	41 days
App only	22 days	31 days	43 days
Web and app	28 days	36 days	56 days
All users	21 days	30 days	42 days

[a]Table indicates at what point in time (days) 75%, 50%, and 25% of users were still using the platform for the different groups.

Figure 1. Nonusage attrition curves for user groups in the subsample of users who were 10,000 Steps members for at least 3 months (n=11,651).



XSL•FO
RenderX

Predictors of Nonusage Attrition

User groups, sociodemographics, engagement, and physical activity as potential predictors of nonusage attrition in the subsample are presented in Table 5. In the univariate analysis, app-only and Web-and-app users showed reduced attrition risk compared to Web-only users (hazard ratio=0.86, 95% CI 0.58-0.68, *P*<.001 and hazard ratio=0.63, 95% CI 0.81-0.93, *P*<.001, respectively). Being a resident of a country other than Australia (hazard ratio=0.87, 95% CI 0.82-0.91, *P*<.001), being male (hazard ratio=0.85, 95% CI 0.82-0.89, *P*<.001), and being of higher age (hazard ratio=0.992, 95% CI 0.991-0.994, *P*<.001) also reduced the risk of attrition. Regarding the engagement parameters, the risk of attrition decreased when the number of individual challenges (hazard ratio=0.62, 95% CI 0.59-0.66, *P*<.001), workplace challenges (hazard ratio=0.94, 95% CI 0.90-0.97, *P*<.001), and number of days physical activity was logged for (hazard ratio=0.921, 95% CI 0.919-0.922, *P*<.001) increased. Furthermore, the more steps logged per day lowered the risk for attrition (hazard ratio=0.99999, 95% CI 0.99998-0.99999, *P*<.001). All variables that had a *P*<.25 in the univariate analysis were included in the multivariate analysis. Results from the multivariate analysis showed a similar pattern (Table 5) except that app-only users did not differ from Web-only users (hazard ratio=0.98, 95% CI 0.91-1.05, *P*=.59) and there was no effect of workplace challenges on attrition risk (hazard ratio=0.98, 95% CI 0.94-1.01, *P*=.19) when controlling for other variables in the analysis.

Table 5. Univariate and multivariate Cox regression: association of nonusage attrition risk with user groups, sociodemographics, engagement, and physical activity in the subsample of users who were 10,000 Steps members for at least 3 months (n=11,651).

Dependent variables	Univariate		Multivariate	
	Hazard ratio (SE), 95% CI	*P*	Hazard ratio (SE), 95% CI	*P*
Group				
Web only	reference		reference	
App only	0.86 (0.03), 0.58-0.68	<.001	0.98 (0.04), 0.91-1.05	.59
Web and app	0.63 (0.03), 0.81-0.93	<.001	0.69 (0.03), 0.64-0.75	<.001
Country				
Australia	reference		reference	
Other	0.87 (0.02), 0.82-0.91	<.001	0.55 (0.02), 0.52-0.58	<.001
Gender				
Female	reference		reference	
Male	0.85 (0.02), 0.82-0.89	<.001	0.95 (0.02), 0.91-0.99	.02
Age	0.992 (0.001), 0.991-0.994	<.001	0.999 (0.00), 0.997-1.000	<.001
Individual challenges	0.62 (0.02), 0.59-0.66	<.001	0.83 (0.03), 0.77-0.89	<.001
Workplace challenges	0.94 (0.02), 0.90-0.97	<.001	0.98 (0.02), 0.94-1.01	.19
Number of days physical activity was logged for	0.921 (0.001), 0.919-0.922	<.001	0.917 (0.001), 0.915-0.918	<.001
Steps per day	0.99999 (0.00000), 0.99998-0.99999	<.001	0.999980 (0.000002), 0.999976-0.999985	<.001

Discussion

Principal Findings

The aim of the present study was to examine program engagement with a freely accessible Internet-delivered physical activity intervention (10,000 Steps Australia) and test for a possible positive effect of using a smartphone app on engagement parameters and attrition. Results indicate a high program engagement and that the use of the app alone or in addition to the website can enhance program engagement and lower attrition. Further, this study extends previous research on individual challenges [20] by showing that workplace challenges were also associated with a prolonged usage of the program (reduced attrition risk).

Program Engagement in Real World Compared to Controlled Settings

The present study reported high levels of program engagement. Nearly all (97%) users logged physical activity at least twice. Whereas most users engaged in at least 1 workplace challenge (81%), only 7% used individual challenges.

Although some studies conducted within controlled settings reported higher program engagement and lower attrition than the 10,000 Steps program [15,35], studies frequently report lower program engagement and higher attrition [36-40]. For example, Funk et al [35] reported a median of 124 exercise logs over 28 months (4.43 logs per months) within their weight loss maintenance program, whereas users from our study with a membership length of at least 9 months (278-297 days) had a median number of physical activity logs of 32 (3.33 logs per months). A study by Steele et al [39] reported 0.98 log-ins per

XSL•FO
RenderX

week over 3 months, whereas our study recorded 3.0 physical activity logs per week for users with a membership length between 2 and 3 months. This is unexpected because all these studies were conducted in controlled settings, whereas our study was not. Studies from controlled trials are generally expected to have better outcomes regarding program usage because participants are likely to be more motivated and committed to the study because of the formal structure and selection process they went through compared to participants from noncontrolled settings [5,6]. Within controlled studies, it may be that the environment where the intervention was delivered plays an important role. Funk et al [35] provided their intervention within primary care clinics; this may have led to a higher commitment to the intervention. In contrast, the other studies recruited participants via local media advertising or email invitations; this provides a less structured environment that may make attrition easier [36-40].

Program Engagement of 10,000 Steps Compared to Other Freely Accessible Programs

There has been some research examining engagement with Web-based interventions in real-life settings including interventions on depression, panic disorder, weight loss, physical activity, drinking behavior, and smoking [15-17,41-46]. Results of this study indicate a high program engagement compared to other freely accessible Web-based interventions. Whereas 97% of users in our study logged physical activity at least 2 times, other freely accessible studies reported between 10% and 62% visiting the intervention at least twice [15,42,43,45,46]. In a previous study on 10,000 Steps, a mean of physical activity logs per week of 1.6 was reported for a study period of 24 months [20]. This is higher than the mean in our study (0.9 logs/week over 9-10 months). This difference is likely caused by the selected sample in that study [20] because participants were users who already used the program for at least 1 month before recruitment and responded to an email invitation. Thus, participants were likely to be more motivated than in our study in which no such selection bias was present. Regarding nonusage attrition, our study showed longer usage of the program compared to other freely accessible interventions. Wanner et al [15] reported a median lifetime usage of 0 days for their physical activity website and after 1 month, only 7% of the registered users were still using the program. Also, Farvolden et al [44] reported only approximately 1% completed their 12-week open-access panic prevention program and Linke et al [41] reported only 24% of users remained in the intervention after 4 weeks for their sensible drinking program. Nevertheless, a commercial Web-based weight loss program [17] showed lower attrition for 12-week subscribers (median lifetime usage of 9 weeks) compared to our data. However, because users had to pay a subscription fee upfront, this have may led to a higher commitment to the intervention compared with studies that are free of any charge [5].

Influence of Smartphone App on Engagement and Nonusage Attrition

The second aim of this study was to examine the effect of a smartphone app on program engagement. In general, app users were younger and more likely to be female compared to Web-only users, which aligns with research showing that the percentage of smartphone owners decreased with age and that young adults are more likely to use health apps [47]. Our results indicate that a smartphone app may assist in prolonging user engagement because using the 10,000 Steps app in addition to the website was associated with a longer duration of usage, higher participation in individual challenges, and higher number of days logging physical activity compared to users who only used the website. Further, nonusage attrition significantly differed across groups; the risk of nonusage attrition was reduced by using the app compared to only using the website to log steps. This effect was more pronounced for Web-and-app users with a 37% smaller attrition hazard compared to app-only users with a 14% smaller attrition hazard when using Web-only users as a reference category. These results are in-line with data from the case-matched control trial of the 10,000 Steps app showing positive effects of the app on the number of steps logged and days physical activity was logged for [27]. Previous research on the influence of smartphone apps on engagement and attrition are scarce. However, for their weight loss promotion app, Carter et al [25] found higher engagement and retention compared to the website diary. In accordance to our study, app users more often logged dietary records compared to the website group (92 days vs 35 days over 6 months). Overall, using smartphones for assistance in health promotion seems appealing because the percentage of people accessing health information via mobile devices is increasing. The Pew mobile health survey [47] found that approximately 52% of smartphone owners reported using their phone to look for health information and 19% have downloaded an app specifically to track or manage health.

Although our results are promising, it should be noted that the majority of users still were Web-only users (83.87%, 14,215/16,948) with only 16.13% (2733/16,948) using the app alone or in addition to the website. These uneven group sizes may be because the app is only available on the iOS mobile platform. Even though rates are increasing, in 2011 the proportion of US adults reporting to own either a smartphone or tablet was 50%, with 38% of smartphone owners and 52% of tablet owners saying their device used the iOS platform [48]. Thus, a substantial number of individuals had to use the website when interested in using the 10,000 Steps program because they either did not have a smartphone or had a mobile device running Android, Blackberry, or Windows. However, we cannot preclude that at least some users chose to use the website over the app intentionally (eg, because of a preference for browser-based surfing or reduced functionality of the app compared to the website).

Sociodemographics, Engagement, and Physical Activity as Determinants of Nonusage Attrition

Personal factors associated with reduced nonusage attrition risk were being male, non-Australian, and older age. Differences in nonusage attrition by country of residency may be seen as an effect of weather because poor and extreme weather has been identified as barriers of physical activity [49] and, in Australia, it is hot and humid for most of the year. Even though evidence shows that females are more likely to be interested in health-related topics (eg, they are more likely to seek online for health information) [50], are more likely to participate in

XSL•FO
RenderX

Web-based physical activity interventions [2,51], and use health apps on their phones more than men [47], in the univariate Cox regression males had a 15% smaller attrition hazard compared to females. However, some research suggests that men are more likely to participate in accelerometer-based studies [52]. Thus, men could be more attracted by technical devices as support for physical activity management (eg, pedometers used in the 10,000 Steps intervention) because they have a more positive attitude toward new technology [53] and, therefore, more interest in maintaining engagement with such an intervention. Effects of age on nonusage attrition are in-line with previous research showing that older age is associated with engagement with the intervention (eg, [45,46]) and as engagement increases, the risk of nonusage attrition decreases. As previous research on the 10,000 Steps project demonstrated [20], the number of individual challenges users participate in is associated with lower attrition risk. Further, this study adds evidence that workplace challenges reduce nonusage attrition risk. This aligns with previous research showing that interactive website components may promote engagement with the intervention [54].

Implications for Future Research

Even though Web-based interventions are capable of reaching large parts of the population, a notable percentage never starts to use or accesses only a small part of the intervention [9]. Because content cannot be helpful if it is not viewed, techniques to enhance engagement with the intervention are needed. Previous research has identified factors that influence exposure and attrition in Internet-delivered interventions [19-22,54]. However, more research is needed to examine effects of such factors in real-life settings. For example, results from Wanner et al [15] suggested that reminder emails are only effective for trial participants, but not for registered open-access users. Previous research has identified peer support as a main facilitator of program engagement [21,54]. This is important because in our study the majority of users heard about the 10,000 Steps program either from their workplace or through a friend. Given the importance of social support as a mediator of behavior change, workplaces especially seem to be a valuable setting for physical activity promotion because its internal structure easily reaches large groups and provides a natural social network [55,56]. The majority of Web-based interventions we compared our results to also used interactive components including discussion boards or goal-setting features [35-38,44], such as the 10,000 Steps program does; however engagement was still higher in our study. This may be due to social support gained through doing workplace challenges within the 10,000 Steps program. This study provided evidence that within real-life settings the use of a smartphone app can enhance engagement with the intervention over time. Most studies report that overall engagement decreases over time (eg, [38]) and high attrition is widely seen as a challenge of Web-based interventions. However, some authors [17,57] argue that this is not necessarily a result of lost interest in the intervention, but of achieving a satisfactory level of behavior change or self-management skills [58]. Future research needs to target reasons for attrition and examine variables in experimental conditions that could distinguish people who decrease and increase engagement over time to design interventions that are likely to be used in long term.

Limitations

A strength of this study is that we obtained a large sample from a freely available physical activity intervention and examined usability efficacy as recommended by previous literature [5]. However, there are some limitations that need to be considered when interpreting the results of this study. First, we only included users with at least 1 physical activity log within the period. Thus, engagement is likely to be higher compared with studies including users with 1 website log-in (people can log in to a website without using any features, such as the 10,000 Steps step log). On the other hand, usage duration was measured as days between first and last physical activity log. This may have underestimated engagement because users could have been active using the discussion board or competing in challenges while not logging steps. However, logging physical activity is the main feature of the 10,000 Steps program and represents a more credible measure than log-ins. Second, our sample included users with varying membership lengths and, therefore, different timeframes of actually being able to use the program. Because usage is likely to decline over time, study length has to be considered when comparing studies. Thus, we reported attrition only for users who were members for at least 3 months. However, we also reported results on engagement data for different membership lengths, which enables comparisons to previous research. Third, in this study we did not report on usage of the discussion forum and virtual walking buddies. This was because app-only users were not able to use these features in the same way as Web-only or Web-and-app users. The discussion forum is not accessible via the app at all; for the virtual walking buddies, users are not able to add buddies via the app. However, previous research has shown that the use of virtual walking buddies was positively associated with the average number of days physical activity was logged for [20]. Lastly, we did not measure physical activity in another form other than steps logged via the program. This is not an objective assessment of participants' activity because logged steps do not necessarily encompass the overall physical activity level of the users. This study did not include an objective measure of physical activity assessing change from preregistration in the program. Although large accelerometer-based studies are emerging (eg, [59]), implementing such measures in the context of this study is challenging due to the timing of assessments. Withstanding their limitations, self-reported data for the period immediately before registering/commencing the program may provide a measure of physical activity that can be used to infer program efficacy in future studies.

Conclusions

Our study provides insight into the engagement with a freely available physical activity intervention. Results indicate that smartphone apps may be powerful tools in enhancing program engagement and lower attrition. Future research should experimentally examine reasons of low engagement and attrition to enable development of interventions that ensure long-term engagement with the intervention. Further, our data elucidate that rating the success of online interventions by engagement

XSL•FO
RenderX

parameters is highly dependent on the time window measured; therefore, study length has to be considered when comparing studies in regard to engagement parameters.

Acknowledgments

Queensland Health provided the funding to Central Queensland University for the development of 10,000 Steps. The development of the iStepLog app was funded by 10,000 Steps. MJD is supported by a Future Leader Fellowship (ID 100029) from the National Heart Foundation of Australia.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Sociodemographics, engagement, and physical activity by user group in the subsample of users who were 10,000 Steps members for at least 3 months (n=11,651).

[PDF File (Adobe PDF File), 151KB - jmir_v17i7e176_app1.pdf]

References

1. World Health Organization. 2014. Global Strategy on Diet, Physical Activity and Health URL: http://www.who.int/dietphysicalactivity/pa/en/ [accessed 2015-02-23] [WebCite Cache ID 6WYhONYbm]
2. Davies CA, Spence JC, Vandelanotte C, Caperchione CM, Mummery WK. Meta-analysis of internet-delivered interventions to increase physical activity levels. Int J Behav Nutr Phys Act 2012;9:52 [FREE Full text] [doi: 10.1186/1479-5868-9-52] [Medline: 22546283]
3. Joseph RP, Durant NH, Benitez TJ, Pekmezi DW. Internet-based physical activity interventions. Am J Lifestyle Med 2014 Dec;8(1):42-68 [FREE Full text] [doi: 10.1177/1559827613498059] [Medline: 25045343]
4. 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]
5. 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]
6. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
7. Motl RW, Dlugonski D, Wójcicki TR, McAuley E, Mohr DC. Internet intervention for increasing physical activity in persons with multiple sclerosis. Mult Scler 2011 Jan;17(1):116-128. [doi: 10.1177/1352458510383148] [Medline: 20921239]
8. Glasgow RE, Christiansen SM, Kurz D, King DK, Woolley T, Faber AJ, et al. Engagement in a diabetes self-management website: usage patterns and generalizability of program use. J Med Internet Res 2011;13(1):e9 [FREE Full text] [doi: 10.2196/jmir.1391] [Medline: 21371992]
9. 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]
10. Norman GJ, Zabinski MF, Adams MA, Rosenberg DE, Yaroch AL, Atienza AA. A review of eHealth interventions for physical activity and dietary behavior change. Am J Prev Med 2007 Oct;33(4):336-345 [FREE Full text] [doi: 10.1016/j.amepre.2007.05.007] [Medline: 17888860]
11. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]
12. Enwald HP, Huotari ML. Preventing the obesity epidemic by second generation tailored health communication: an interdisciplinary review. J Med Internet Res 2010;12(2):e24 [FREE Full text] [doi: 10.2196/jmir.1409] [Medline: 20584698]
13. Christensen H, Griffiths KM, Farrer L. Adherence in Internet interventions for anxiety and depression. J Med Internet Res 2009;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]
14. 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]
15. Wanner M, Martin-Diener E, Bauer G, Braun-Fahrländer C, Martin BW. Comparison of trial participants and open access users of a web-based physical activity intervention regarding adherence, attrition, and repeated participation. J Med Internet Res 2010;12(1):e3 [FREE Full text] [doi: 10.2196/jmir.1361] [Medline: 20147006]

XSL•FO
RenderX

16. Christensen H, Griffiths KM, Korten AE, Brittliffe K, Groves C. A comparison of changes in anxiety and depression symptoms of spontaneous users and trial participants of a cognitive behavior therapy website. J Med Internet Res 2004 Dec 22;6(4):e46 [FREE Full text] [doi: 10.2196/jmir.6.4.e46] [Medline: 15631970]
17. Neve MJ, Collins CE, Morgan PJ. Dropout, nonusage attrition, and pretreatment predictors of nonusage attrition in a commercial Web-based weight loss program. J Med Internet Res 2010;12(4):e69 [FREE Full text] [doi: 10.2196/jmir.1640] [Medline: 21156470]
18. Hutchesson MJ, Collins CE, Morgan PJ, Callister R. An 8-week web-based weight loss challenge with celebrity endorsement and enhanced social support: observational study. J Med Internet Res 2013;15(7):e129 [FREE Full text] [doi: 10.2196/jmir.2540] [Medline: 23827796]
19. Schneider F, de Vries H, Candel M, van de Kar A, van Osch L. Periodic email prompts to re-use an internet-delivered computer-tailored lifestyle program: influence of prompt content and timing. J Med Internet Res 2013;15(1):e23 [FREE Full text] [doi: 10.2196/jmir.2151] [Medline: 23363466]
20. Davies C, Corry K, Van Itallie A, Vandelanotte C, Caperchione C, Mummery WK. Prospective associations between intervention components and website engagement in a publicly available physical activity website: the case of 10,000 Steps Australia. J Med Internet Res 2012;14(1):e4 [FREE Full text] [doi: 10.2196/jmir.1792] [Medline: 22260810]
21. Richardson CR, Buis LR, Janney AW, Goodrich DE, Sen A, Hess ML, et al. An online community improves adherence in an internet-mediated walking program. Part 1: results of a randomized controlled trial. J Med Internet Res 2010;12(4):e71 [FREE Full text] [doi: 10.2196/jmir.1338] [Medline: 21169160]
22. Lewis B, Williams D, Dunsiger S, Sciamanna C, Whiteley J, Napolitano M, et al. User attitudes towards physical activity websites in a randomized controlled trial. Prev Med 2008 Nov;47(5):508-513 [FREE Full text] [doi: 10.1016/j.ypmed.2008.07.020] [Medline: 18773915]
23. Brouwer W, Oenema A, Raat H, Crutzen R, de Nooijer J, de Vries NK, et al. Characteristics of visitors and revisitors to an Internet-delivered computer-tailored lifestyle intervention implemented for use by the general public. Health Educ Res 2010 Aug;25(4):585-595 [FREE Full text] [doi: 10.1093/her/cyp063] [Medline: 19897515]
24. Bort-Roig J, Gilson ND, Puig-Ribera A, Contreras RS, Trost SG. Measuring and influencing physical activity with smartphone technology: a systematic review. Sports Med 2014 May;44(5):671-686. [doi: 10.1007/s40279-014-0142-5] [Medline: 24497157]
25. Carter MC, Burley VJ, Nykjaer C, Cade JE. Adherence to a smartphone application for weight loss compared to website and paper diary: pilot randomized controlled trial. J Med Internet Res 2013;15(4):e32 [FREE Full text] [doi: 10.2196/jmir.2283] [Medline: 23587561]
26. Cobb NK, Poirier J. Effectiveness of a multimodal online well-being intervention: a randomized controlled trial. Am J Prev Med 2014 Jan;46(1):41-48 [FREE Full text] [doi: 10.1016/j.amepre.2013.08.018] [Medline: 24355670]
27. Kirwan M, Duncan MJ, Vandelanotte C, Mummery WK. Using smartphone technology to monitor physical activity in the 10,000 Steps program: a matched case-control trial. J Med Internet Res 2012;14(2):e55 [FREE Full text] [doi: 10.2196/jmir.1950] [Medline: 22522112]
28. Brown WJ, Mummery K, Eakin EG, Schofield G. 10,000 Steps Rockhampton: Evaluation of a whole community approach to improving population levels of physical activity. J Phys Act Health 2006;3(1):1-14 [FREE Full text]
29. Antypas K, Wangberg SC. An Internet- and mobile-based tailored intervention to enhance maintenance of physical activity after cardiac rehabilitation: short-term results of a randomized controlled trial. J Med Internet Res 2014;16(3):e77 [FREE Full text] [doi: 10.2196/jmir.3132] [Medline: 24618349]
30. Tudor-Locke C, Ham SA, Macera CA, Ainsworth BE, Kirtland KA, Reis JP, et al. Descriptive epidemiology of pedometer-determined physical activity. Med Sci Sports Exerc 2004 Sep;36(9):1567-1573. [Medline: 15354039]
31. De Cocker KA, De Bourdeaudhuij IM, Brown WJ, Cardon GM. Effects of "10,000 steps Ghent": a whole-community intervention. Am J Prev Med 2007 Dec;33(6):455-463. [doi: 10.1016/j.amepre.2007.07.037] [Medline: 18022061]
32. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Hillsdale, NJ: L Erlbaum Associates; 1988.
33. 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 Dec;10(4):e56 [FREE Full text] [doi: 10.2196/jmir.1108] [Medline: 19117828]
34. Hosmer DW, Lemeshow S, Sturdivant RX. Model-building strategies and methods for logistic regression. In: Applied Logistic Regression. New York: Wiley; 2013:89-153.
35. Funk KL, Stevens VJ, Appel LJ, Bauck A, Brantley PJ, Champagne CM, et al. Associations of Internet website use with weight change in a long-term weight loss maintenance program. J Med Internet Res 2010;12(3):e29 [FREE Full text] [doi: 10.2196/jmir.1504] [Medline: 20663751]
36. Danaher BG, Boles SM, Akers L, Gordon JS, Severson HH. Defining participant exposure measures in Web-based health behavior change programs. J Med Internet Res 2006;8(3):e15 [FREE Full text] [doi: 10.2196/jmir.8.3.e15] [Medline: 16954125]
37. Brindal E, Freyne J, Saunders I, Berkovsky S, Smith G, Noakes M. Features predicting weight loss in overweight or obese participants in a web-based intervention: randomized trial. J Med Internet Res 2012;14(6):e173 [FREE Full text] [doi: 10.2196/jmir.2156] [Medline: 23234759]

XSL•FO
RenderX

38. McKay HG, King D, Eakin EG, Seeley JR, Glasgow RE. The diabetes network internet-based physical activity intervention: a randomized pilot study. Diabetes Care 2001 Aug;24(8):1328-1334. [Medline: 11473065]
39. Steele R, Mummery WK, Dwyer T. Using the Internet to promote physical activity: a randomized trial of intervention delivery modes. J Phys Act Health 2007 Jul;4(3):245-260. [Medline: 17846455]
40. Hurling R, Catt M, Boni MD, Fairley BW, Hurst T, Murray P, et al. Using Internet and mobile phone technology to deliver an automated physical activity program: randomized controlled trial. J Med Internet Res 2007;9(2):e7 [FREE Full text] [doi: 10.2196/jmir.9.2.e7] [Medline: 17478409]
41. Linke S, Murray E, Butler C, Wallace P. Internet-based interactive health intervention for the promotion of sensible drinking: patterns of use and potential impact on members of the general public. J Med Internet Res 2007;9(2):e10 [FREE Full text] [doi: 10.2196/jmir.9.2.e10] [Medline: 17513281]
42. Wang J, Etter J. Administering an effective health intervention for smoking cessation online: the international users of Stop-Tabac. Prev Med 2004 Nov;39(5):962-968. [doi: 10.1016/j.ypmed.2004.03.038] [Medline: 15475030]
43. Balmford J, Borland R, Benda P. Patterns of use of an automated interactive personalized coaching program for smoking cessation. J Med Internet Res 2008;10(5):e54 [FREE Full text] [doi: 10.2196/jmir.1016] [Medline: 19097975]
44. Farvolden P, Denisoff E, Selby P, Bagby RM, Rudy L. Usage and longitudinal effectiveness of a Web-based self-help cognitive behavioral therapy program for panic disorder. J Med Internet Res 2005;7(1):e7 [FREE Full text] [doi: 10.2196/jmir.7.1.e7] [Medline: 15829479]
45. Lara MA, Tiburcio M, Aguilar AA, Sánchez-Solís A. A four-year experience with a Web-based self-help intervention for depressive symptoms in Mexico. Rev Panam Salud Publica 2014;35(5-6):399-406 [FREE Full text] [Medline: 25211568]
46. Verheijden MW, Jans MP, Hildebrandt VH, Hopman-Rock M. Rates and determinants of repeated participation in a web-based behavior change program for healthy body weight and healthy lifestyle. J Med Internet Res 2007;9(1):e1 [FREE Full text] [doi: 10.2196/jmir.9.1.e1] [Medline: 17478410]
47. Fox S, Duggan M. Mobile Health 2012. Washington, DC: Pew Internet & American Life Project; 2012 Nov 08. URL: http://www.pewinternet.org/files/old-media//Files/Reports/2012/PIP_MobileHealth2012_FINAL.pdf [accessed 2015-02-23] [WebCite Cache ID 6WYhei59e]
48. The Future of Mobile News: The Explosion in Mobile Audiences and a Close Look at What it Means for News. Washington, DC: Pew Internet & American Life Project; 2012 Oct 01. URL: http://www.journalism.org/files/legacy/Futureofmobilenews%20_final1.pdf [accessed 2015-04-01] [WebCite Cache ID 6XSd5y57k]
49. 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]
50. Fox S, Duggan M. Health Online 2013. Washington, DC: Pew Internet & American Life Project; 2013 Jan 15. URL: http://www.pewinternet.org/files/old-media//Files/Reports/PIP_HealthOnline.pdf [accessed 2015-05-18] [WebCite Cache ID 6YcFx73VK]
51. Vandelanotte C, Spathonis KM, Eakin EG, Owen N. Website-delivered physical activity interventions a review of the literature. Am J Prev Med 2007 Jul;33(1):54-64. [doi: 10.1016/j.amepre.2007.02.041] [Medline: 17572313]
52. Harris TJ, Victor CR, Carey IM, Adams R, Cook DG. Less healthy, but more active: opposing selection biases when recruiting older people to a physical activity study through primary care. BMC Public Health 2008;8:182 [FREE Full text] [doi: 10.1186/1471-2458-8-182] [Medline: 18505574]
53. Schumacher P, Morahan-Martin J. Gender, Internet and computer attitudes and experiences. Computers in Human Behavior 2001 Jan;17(1):95-110. [doi: 10.1016/S0747-5632(00)00032-7]
54. Brouwer W, Kroeze W, Crutzen R, de Nooijer J, de Vries NK, Brug J, et al. Which intervention characteristics are related to more exposure to internet-delivered healthy lifestyle promotion interventions? A systematic review. J Med Internet Res 2011;13(1):e2 [FREE Full text] [doi: 10.2196/jmir.1639] [Medline: 21212045]
55. Lauzon N, Chan CB, Myers AM, Tudor-Locke C. Participant experiences in a workplace pedometer-based physical activity program. J Phys Act Health 2008 Sep;5(5):675-687. [Medline: 18820340]
56. Robroek SJ, van Lenthe FJ, van Empelen P, Burdorf A. Determinants of participation in worksite health promotion programmes: a systematic review. Int J Behav Nutr Phys Act 2009;6:26 [FREE Full text] [doi: 10.1186/1479-5868-6-26] [Medline: 19457246]
57. 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]
58. Bandura A. The primacy of self-regulation in health promotion. Applied Psychology 2005 Apr;54(2):245-254. [doi: 10.1111/j.1464-0597.2005.00208.x]
59. Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc 2008 Jan;40(1):181-188. [doi: 10.1249/mss.0b013e31815a51b3] [Medline: 18091006]

XSL•FO
RenderX

Edited by G Eysenbach; submitted 03.03.15; peer-reviewed by A Burdorf, A Middelweerd, J Brewer, K Woods-Townsend; comments to author 25.03.15; revised version received 08.06.15; accepted 29.06.15; published 15.07.15.

Please cite as:
Guertler D, Vandelanotte C, Kirwan M, Duncan MJ
Engagement and Nonusage Attrition With a Free Physical Activity Promotion Program: The Case of 10,000 Steps Australia
J Med Internet Res 2015;17(7):e176
URL: http://www.jmir.org/2015/7/e176/
doi:10.2196/jmir.4339
PMID:26180040

©Diana Guertler, Corneel Vandelanotte, Morwenna Kirwan, Mitch J Duncan. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 15.07.2015. 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 Baseline Assessment Modality on Enrollment and Retention in a Facebook Smoking Cessation Study

Andrea C Villanti[1,2], MPH, PhD; Megan A Jacobs[1], MPH; Grace Zawistowski[1,3], BA; Jody Brookover[1], BS; Cassandra A Stanton[4,5], PhD; Amanda L Graham[1,4], PhD

[1]The Schroeder Institute for Tobacco Research and Policy Studies, Legacy, Washington, DC, United States

[2]Department of Health, Behavior and Society, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

[3]Milken Institute School of Public Health, The George Washington University, Washington, DC, United States

[4]Department of Oncology, Georgetown University Medical Center/Cancer Prevention and Control Program, Lombardi Comprehensive Cancer Center, Washington, DC, United States

[5]Westat, Rockville, MD, United States

Corresponding Author:
Andrea C Villanti, MPH, PhD
The Schroeder Institute for Tobacco Research and Policy Studies
Legacy
1724 Massachusetts Ave NW
Washington, DC, 20036
United States
Phone: 1 202 454 5751
Fax: 1 202 454 5785
Email: avillanti@legacyforhealth.org

Abstract

Background: Few studies have addressed enrollment and retention methods in online smoking cessation interventions. Fully automated Web-based trials can yield large numbers of participants rapidly but suffer from high rates of attrition. Personal contact with participants can increase recruitment of smokers into cessation trials and improve participant retention.

Objective: To compare the impact of Web-based (WEB) and phone (PH) baseline assessments on enrollment and retention metrics in the context of a Facebook smoking cessation study.

Methods: Participants were recruited via Facebook and Google ads which were randomly displayed to adult smokers in the United States over 27 days from August to September 2013. On each platform, two identical ads were randomly displayed to users who fit the advertising parameters. Clicking on one of the ads resulted in randomization to WEB, and clicking on the other ad resulted in randomization to PH. Following online eligibility screening and informed consent, participants in the WEB arm completed the baseline survey online whereas PH participants completed the baseline survey by phone with a research assistant. All participants were contacted at 30 days to complete a follow-up survey that assessed use of the cessation intervention and smoking outcomes. Participants were paid $15 for follow-up survey completion.

Results: A total of 4445 people clicked on the WEB ad and 4001 clicked on the PH ad: 12.04% (n=535) of WEB participants and 8.30% (n=332) of PH participants accepted the online study invitation (P<.001). Among the 726 participants who completed online eligibility screening, an equivalent proportion in both arms was eligible and an equivalent proportion of the eligible participants in both arms provided informed consent. There was significant drop-off between consent and completion of the baseline survey in the PH arm, resulting in enrollment rates of 32.7% (35/107) for the PH arm and 67.9% (114/168) for the WEB arm (P<.001). The overall enrollment rate among everyone who clicked on a study ad was 2%. There were no between group differences in the proportion that installed the Facebook app (66/114, 57.9% WEB vs 17/35, 49% PH) or that completed the 30-day follow-up survey (49/114, 43.0% WEB vs 16/35, 46% PH). A total of $6074 was spent on ads, generating 3,834,289 impressions and resulting in 8446 clicks (average cost $0.72 per click). Per participant enrollment costs for advertising alone were $27 WEB and $87 PH.

Conclusions: A more intensive phone baseline assessment protocol yielded a lower rate of enrollment, equivalent follow-up rates, and higher enrollment costs compared to a Web-based assessment protocol. Future research should focus on honing mixed-mode assessment protocols to further optimize enrollment and retention.

XSL•FO
RenderX

(J Med Internet Res 2015;17(7):e179) doi:10.2196/jmir.4341

KEYWORDS

research subject recruitment; smoking cessation; Internet; social networking; adult

Introduction

Systematic reviews of recruitment methods for health studies [1-4] have highlighted online recruitment as a way to address many central enrollment challenges. Advantages of online recruitment include the ability to rapidly and cost-effectively reach a broad population, including those typically defined as hard to reach [5,6]. However, low retention (ie, loss to follow-up [7]) is of particular concern in fully automated, online trials where follow-up rates are often lower than in-person trials [8-10]. Internet-based trials have yielded an average follow-up rate of 53% for fully automated randomized trials, with some studies reporting rates as low as 13% [8].

Several approaches have been suggested to improve retention in Internet-based trials, including offline consent and data collection [10-13], modified online data collection formats [9,14-16], and filtering potential participants based on characteristics correlated with higher response rates [9,17-21]. There is evidence that online trials employing both online and offline follow-up methods may yield higher rates of follow-up [8,12,22,23].

This study used a randomized design to compare the effect of Internet and telephone baseline assessments on recruitment and retention metrics in a smoking cessation study involving Facebook. Our a priori hypothesis was that a more personalized assessment strategy at baseline (by phone) would depress enrollment rates and result in greater enrollment costs but would increase retention compared to a fully automated baseline assessment via the Internet. The study was conducted from August 2013 through November 2013 and was approved by the Schulman Associates Institutional Review Board.

Methods

Participants

Eligible participants had to be adult (18 years or older), self-identified smokers who were thinking of quitting in the next 30 days, had an active Facebook account, and had not used the smoking cessation Facebook app UbiQUITous [24].

Recruitment, Enrollment, and Randomization

Facebook and Google AdWords advertisements were implemented simultaneously for 27 days from August 28 through September 24, 2013. All ads targeted adult smokers within the United States and were designed to be as similar as possible given the differences between Facebook and Google advertising platforms. On each platform, two identical ads were randomly displayed to users who fit the advertising parameters. Clicking on one of the ads resulted in randomization to the Internet baseline assessment (WEB), and clicking on the other ad resulted in randomization to the phone baseline assessment (PH).

A daily spending limit was set at USD $300 for all advertising. The bidding structure was set to maximize impressions in Facebook [25] and to maximize clicks in Google, with a max bid of $0.25/click. This advertising approach built on lessons learned by our research group from a previous randomized trial conducted within Facebook [24] and was consistent with each platform's best practices at the time of the study [25-27].

Clicking on any of the ads took users to a study invitation page that provided a brief overview of the study and invited individuals to complete eligibility screening. Eligibility screening assessed gender, ethnicity, race, education, smoking status, motivation to quit, age, whether they had an active Facebook account, and whether they had used the Facebook app involved in this study. Only the last 5 questions determined eligibility; the other information was used to examine the characteristics of individuals responding to the recruitment ads.

Eligible participants completed online informed consent and provided contact information. Participants randomized to the WEB condition were immediately directed to the online baseline assessment; participants in the PH condition were informed that they would be contacted within 48 hours by a research assistant to complete the baseline assessment.

Intervention

Following completion of the baseline assessment, all participants received information on how to install the study intervention, UbiQUITous, a Facebook app grounded in evidence-based smoking cessation treatment guidelines [28]. The development and implementation of the app is described in detail elsewhere [24].

Assessment Procedures

The baseline survey comprised 14 questions addressing tobacco use, cessation-related cognitions and behaviors, and intensity of Facebook use [29]. It was deliberately short so as not to represent a barrier to study enrollment. At 30 days post-enrollment, participants in both arms were asked to complete a brief follow-up survey. Survey invitations were delivered via email, Facebook message, or telephone, with telephone follow-up for nonresponders. The survey assessed point prevalence abstinence (30-day and 7-day), tobacco use among those not abstinent, cessation intentions and quit methods used, and satisfaction with the UbiQUITous app. Participants were paid $15 for completion of the follow-up survey (Amazon gift certificate delivered via email).

Outcomes

The outcomes of interest were the proportion enrolling in the study in the WEB compared to the PH arm and of those enrolled, the proportion completing the follow-up survey at 30 days (retention) in the WEB compared to the PH condition. Secondary outcomes were recruitment volume and enrollment costs.

XSL•FO
RenderX

Statistical Analysis

This study used two sources of data: advertising metrics and enrollment and follow-up data extracted from our clinical trials management system. First, we developed a CONSORT diagram to track participants from ad exposure through enrollment and retention and estimated differences in proportions at each stage using chi-square tests. Second, we estimated the advertising cost per randomized participant in the WEB and PH conditions.

Results

Enrollment Metrics

Table 1 presents the flow of participants from online recruitment through follow-up. Overall, 4445 people were referred to the study from the WEB ad and 4001 from the PH ad; 12.04% (535/4445) of WEB participants and 8.30% (332/4001) of PH participants accepted the online study invitation (P<.001). There were no between-group differences in rates of eligibility or informed consent. The main reason for ineligibility was lack of intention to quit within the next 30 days (125 WEB, 57 PH). There was significant drop-off between consent and completion of the baseline survey in the PH arm, resulting in an enrollment rate of 32.7% (35/107) of PH participants versus 67.9% (114/168) of WEB participants (P<.001). The overall enrollment rate among everyone who clicked on a study ad was 1.76% (149/8446). Of the 149 participants who were enrolled, only 83 (55.7%) installed the Facebook app, with no differences between study arms (WEB: 66/114, 57.9% vs PH: 17/35, 49%; P=.33).

Retention Metrics

As shown in Table 1, 43.6% of enrolled participants completed the 30-day follow-up survey, with no differences between study arms (WEB: 49/114, 43.0% vs PH: 16/35, 46%; P=.78).

Enrollment Costs

During the recruitment period, a total of $6074 (WEB $3027; PH $3047) was spent on online recruitment ads, generating 3,834,289 (WEB 2,030,253; PH 1,804,036) impressions and resulting in 8446 (WEB 4445; PH 4001) participants clicking on the ads. Advertising costs per randomized participant were $27 WEB and $87 PH.

Table 1. Enrollment flow of participants.

Enrollment Step	Total, n (%)	WEB Arm, n (%)	PH Arm, n (%)	*P* value
Clicked on recruitment ad	8446 (100%)	4445 (100%)	4001 (100%)	—
Accepted study invite	867/8446 (10.3%)	535/4445 (12.0%)	332/4001 (8.3%)	<.001
Completed eligibility screening	726/867 (83.7%)	461/535 (86.2%)	265/332 (79.8%)	.014
Eligible	474/726 (65.3%)	292/461 (63.3%)	182/265 (68.7%)	.15
Consented	275/474 (58.0%)	168/292 (57.5%)	107/182 (58.8%)	.79
Enrolled	149/275 (54.2%)	114/168 (67.9%)	35/107 (32.7%)	<.001
Installed Facebook app	83/149 (55.7%)	66/114 (57.9%)	17/35 (48.6%)	.33
Completed follow-up	65/149 (43.6%)	49/114 (43.0%)	16/35 (45.7%)	.78

Discussion

Principal Findings

This study used a novel randomized design to assess the impact of baseline assessment method on recruitment and retention. Our main finding was that a more intensive phone baseline assessment protocol yielded a lower rate of enrollment, equivalent follow-up rates, and higher enrollment costs compared to a web-based baseline assessment protocol. The overall enrollment rate of 2% was smaller than other trials but still in the range of 6% observed in another large web-based randomized trial [13]. Overall retention at 30 days in our study was 44%, similar to previous literature on web-based randomized trials [8] and cohort studies [30] but lower than reported in other web-based cessation studies with longer follow-up periods [13,31]. The equivalent retention rates in the web and phone arms are consistent with findings reported in a weight management study [32]. Our enrollment costs were also in line with recent studies recruiting via Facebook ads [26,27].

Strengths and Limitations

While our study suffered from low enrollment, differential drop-out between study arms, and low retention, these metrics were in line with those reported in other online trials [8,13]. Of interest to others conducting online recruitment, we found Google AdWords to be an ineffective recruitment strategy driven by automated shut-off by Google due to poor performance. Possible barriers to performance from the Google ads may have been the frame shift needed from an Internet search query to Facebook app installation (versus Facebook ad for a Facebook app) or the design of the ads, which attempted to keep content and pricing similar in order to directly compare the yield of Facebook to Google advertising. Researchers using online ads across different platforms for study recruitment should take these concerns into consideration when designing and implementing recruitment protocols.

One strength of our study is the randomization of participants to recruitment method using the underlying auction mechanism to place ads within Facebook and Google. Our study is one of the first to explicitly examine the impact of enrollment strategies on recruitment and retention metrics within the context of online intervention research. Future research should focus on honing advertising strategies and web-based assessment protocols to further optimize enrollment and retention.

XSL•FO
RenderX

Acknowledgments

This study was funded by the American Legacy Foundation. We would like to thank Eric Palakovich Carr and Danae Johnson for their contributions to this project. M. Jacobs and A. Graham were supported in part by R01 CA155369-01A1. C. Stanton was supported in part by the National Cancer Institute award P30CA051008.

Conflicts of Interest

None declared.

References

1. Treweek S, Pitkethly M, Cook J, Kjeldstrøm M, Taskila T, Johansen M, et al. Strategies to improve recruitment to randomised controlled trials. Cochrane Database Syst Rev 2010(4):MR000013. [doi: 10.1002/14651858.MR000013.pub5] [Medline: 20393971]
2. Caldwell PH, Hamilton S, Tan A, Craig JC. Strategies for increasing recruitment to randomised controlled trials: systematic review. PLoS Med 2010;7(11):e1000368 [FREE Full text] [doi: 10.1371/journal.pmed.1000368] [Medline: 21085696]
3. Marcano Belisario JS, Bruggeling MN, Gunn LH, Brusamento S, Car J. Interventions for recruiting smokers into cessation programmes. Cochrane Database Syst Rev 2012;12:CD009187. [doi: 10.1002/14651858.CD009187.pub2] [Medline: 23235672]
4. Mapstone J, Elbourne D, Roberts I. Strategies to improve recruitment to research studies. Cochrane Database Syst Rev 2007(2):MR000013. [doi: 10.1002/14651858.MR000013.pub3] [Medline: 17443634]
5. Rhodes SD, Bowie DA, Hergenrather KC. Collecting behavioural data using the world wide web: considerations for researchers. J Epidemiol Community Health 2003 Jan;57(1):68-73 [FREE Full text] [Medline: 12490652]
6. Graham AL, Milner P, Saul JE, Pfaff L. Online advertising as a public health and recruitment tool: comparison of different media campaigns to increase demand for smoking cessation interventions. J Med Internet Res 2008;10(5):e50 [FREE Full text] [doi: 10.2196/jmir.1001] [Medline: 19073542]
7. 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]
8. Mathieu E, McGeechan K, Barratt A, Herbert R. Internet-based randomized controlled trials: a systematic review. J Am Med Inform Assoc 2013 May 1;20(3):568-576 [FREE Full text] [doi: 10.1136/amiajnl-2012-001175] [Medline: 23065196]
9. Murray E, White IR, Varagunam M, Godfrey C, Khadjesari Z, McCambridge J. Attrition revisited: adherence and retention in a web-based alcohol trial. J Med Internet Res 2013;15(8):e162 [FREE Full text] [doi: 10.2196/jmir.2336] [Medline: 23996958]
10. Shorter GW, Ferry F. Recruitment and retention in internet based randomised trials. Trials 2013;14(Suppl 1):O113. [doi: 10.1186/1745-6215-14-S1-O113]
11. Bull SS, Lloyd L, Rietmeijer C, McFarlane M. Recruitment and retention of an online sample for an HIV prevention intervention targeting men who have sex with men: the Smart Sex Quest Project. AIDS Care 2004 Nov;16(8):931-943. [doi: 10.1080/09540120412331292507] [Medline: 15511725]
12. Bull SS, Vallejos D, Levine D, Ortiz C. Improving recruitment and retention for an online randomized controlled trial: experience from the Youthnet study. AIDS Care 2008 Sep;20(8):887-893. [doi: 10.1080/09540120701771697] [Medline: 18777217]
13. Graham AL, Cobb NK, Papandonatos GD, Moreno JL, Kang H, Tinkelman DG, et al. A randomized trial of Internet and telephone treatment for smoking cessation. Arch Intern Med 2011 Jan 10;171(1):46-53 [FREE Full text] [doi: 10.1001/archinternmed.2010.451] [Medline: 21220660]
14. McCambridge J, Kalaitzaki E, White IR, Khadjesari Z, Murray E, Linke S, et al. Impact of length or relevance of questionnaires on attrition in online trials: randomized controlled trial. J Med Internet Res 2011;13(4):e96 [FREE Full text] [doi: 10.2196/jmir.1733] [Medline: 22100793]
15. Ekman A, Klint A, Dickman P, Adami H, Litton J. Optimizing the design of web-based questionnaires--experience from a population-based study among 50,000 women. Eur J Epidemiol 2007;22(5):293-300. [doi: 10.1007/s10654-006-9091-0] [Medline: 17206467]
16. Edwards P. Questionnaires in clinical trials: guidelines for optimal design and administration. Trials 2010;11:2 [FREE Full text] [doi: 10.1186/1745-6215-11-2] [Medline: 20064225]
17. Mathew M, Morrow JR, Frierson GM, Bain TM. Assessing digital literacy in web-based physical activity surveillance: the WIN study. Am J Health Promot 2011;26(2):90-95. [doi: 10.4278/ajhp.091001-QUAN-320] [Medline: 22040389]
18. Khosropour CM, Sullivan PS. Predictors of retention in an online follow-up study of men who have sex with men. J Med Internet Res 2011;13(3):e47 [FREE Full text] [doi: 10.2196/jmir.1717] [Medline: 21745792]

XSL•FO
RenderX

19. Sullivan PS, Khosropour CM, Luisi N, Amsden M, Coggia T, Wingood GM, et al. Bias in online recruitment and retention of racial and ethnic minority men who have sex with men. J Med Internet Res 2011;13(2):e38 [FREE Full text] [doi: 10.2196/jmir.1797] [Medline: 21571632]
20. Glasgow RE, Nelson CC, Kearney KA, Reid R, Ritzwoller DP, Strecher VJ, et al. Reach, engagement, and retention in an Internet-based weight loss program in a multi-site randomized controlled trial. J Med Internet Res 2007;9(2):e11 [FREE Full text] [doi: 10.2196/jmir.9.2.e11] [Medline: 17513282]
21. Verheijden MW, Jans MP, Hildebrandt VH, Hopman-Rock M. Rates and determinants of repeated participation in a web-based behavior change program for healthy body weight and healthy lifestyle. J Med Internet Res 2007;9(1):e1 [FREE Full text] [doi: 10.2196/jmir.9.1.e1] [Medline: 17478410]
22. Sackett DL. The principles behind the tactics of performing therapeutic trials. In: Haynes RB, Sackett DL, Guyatt GH, Tugwell P, editors. Clinical epidemiology: how to do clinical practice research. Philadelphia: Lippincott Williams & Wilkins; 2006:173-243.
23. Thorpe KE, Zwarenstein M, Oxman AD, Treweek S, Furberg CD, Altman DG, et al. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. J Clin Epidemiol 2009 May;62(5):464-475. [doi: 10.1016/j.jclinepi.2008.12.011] [Medline: 19348971]
24. Cobb NK, Jacobs MA, Saul J, Wileyto EP, Graham AL. Diffusion of an evidence-based smoking cessation intervention through Facebook: a randomised controlled trial study protocol. BMJ Open 2014;4(1):e004089 [FREE Full text] [doi: 10.1136/bmjopen-2013-004089] [Medline: 24448847]
25. Optimized CPM.: Facebook Inc URL: https://developers.facebook.com/docs/reference/ads-api/optimizedcpm [accessed 2015-02-05]
26. Ramo DE, Hall SM, Prochaska JJ. Reaching young adult smokers through the Internet: comparison of three recruitment mechanisms. Nicotine Tob Res 2010 Jul;12(7):768-775 [FREE Full text] [doi: 10.1093/ntr/ntq086] [Medline: 20530194]
27. Ramo DE, Rodriguez TMS, Chavez K, Sommer MJ, Prochaska JJ. Facebook Recruitment of Young Adult Smokers for a Cessation Trial: Methods, Metrics, and Lessons Learned. Internet Interv 2014 Apr;1(2):58-64 [FREE Full text] [doi: 10.1016/j.invent.2014.05.001] [Medline: 25045624]
28. The Tobacco Use and Dependence Clinical Practice Guideline Panel, Staff, and Consortium Representatives. A clinical practice guideline for treating tobacco use and dependence: A US Public Health Service report. JAMA 2000 Jun 28;283(24):3244-3254. [Medline: 10866874]
29. Ellison NB, Steinfield C, Lampe C. The Benefits of Facebook "Friends:" Social Capital and College Students' Use of Online Social Network Sites. J Comput Mediat Commun 2007;12(4):1143-1168.
30. Bajardi P, Paolotti D, Vespignani A, Eames K, Funk S, Edmunds WJ, et al. Association between recruitment methods and attrition in Internet-based studies. PLoS One 2014;9(12):e114925 [FREE Full text] [doi: 10.1371/journal.pone.0114925] [Medline: 25490045]
31. Richardson A, Graham AL, Cobb N, Xiao H, Mushro A, Abrams D, et al. Engagement promotes abstinence in a web-based cessation intervention: cohort study. J Med Internet Res 2013;15(1):e14 [FREE Full text] [doi: 10.2196/jmir.2277] [Medline: 23353649]
32. Funk KL, Elder CR, Lindberg NM, Gullion CM, DeBar LL, Meltesen G, et al. Comparison of characteristics and outcomes by initial study contact (website versus staff) for participants enrolled in a weight management study. Clin Trials 2012 Apr;9(2):226-231 [FREE Full text] [doi: 10.1177/1740774511432727] [Medline: 22273589]

Edited by G Eysenbach; submitted 11.02.15; peer-reviewed by D Ramo, L Struik; comments to author 26.03.15; revised version received 11.04.15; accepted 28.04.15; published 16.07.15.

Please cite as:
Villanti AC, Jacobs MA, Zawistowski G, Brookover J, Stanton CA, Graham AL
Impact of Baseline Assessment Modality on Enrollment and Retention in a Facebook Smoking Cessation Study
J Med Internet Res 2015;17(7):e179
URL: http://www.jmir.org/2015/7/e179/
doi:10.2196/jmir.4341
PMID:26183789

©Andrea C Villanti, Megan A Jacobs, Grace Zawistowski, Jody Brookover, Cassandra A Stanton, Amanda L Graham. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 16.07.2015. 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

Effect of a Web-Based Behavior Change Program on Weight Loss and Cardiovascular Risk Factors in Overweight and Obese Adults at High Risk of Developing Cardiovascular Disease: Randomized Controlled Trial

Sinead Watson[1,2], BSc, MSc, PhD; Jayne V Woodside[1], MA, PhD; Lisa J Ware[3], BSc, MSc, PhD; Steven J Hunter[4], MB, BCh, MD, FRCP; Alanna McGrath[1], BSc, PhD; Christopher R Cardwell[1], BSc, MSc, PhD; Katherine M Appleton[5], BSc, PhD; Ian S Young[1], BSc, MB, BCh, MD, FRCP; Michelle C McKinley[1], BSc, PhD

[1]Centre for Public Health, School of Medicine, Dentistry and Biomedical Sciences, Queen's University Belfast, Belfast, United Kingdom

[2]Institute for Global Food Security, School of Biological Sciences, Queen's University Belfast, Belfast, United Kingdom

[3]Hypertension in Africa Research Team (HART), Faculty of Health Sciences, North-West University, Potchefstroom, South Africa

[4]Regional Centre for Endocrinology and Diabetes, Royal Victoria Hospital, Belfast, United Kingdom

[5]Department of Psychology, Faculty of Science and Technology, Bournemouth University, Poole, United Kingdom

Corresponding Author:
Michelle C McKinley, BSc, PhD
Centre for Public Health
School of Medicine, Dentistry and Biomedical Sciences
Queen's University Belfast
Institute of Clinical Science Block B
Grosvenor Road
Belfast, BT12 6BJ
United Kingdom
Phone: 44 (0)28 9063 2685
Fax: 44 (0)28 90235900
Email: m.mckinley@qub.ac.uk

Abstract

Background: Web-based programs are a potential medium for supporting weight loss because of their accessibility and wide reach. Research is warranted to determine the shorter- and longer-term effects of these programs in relation to weight loss and other health outcomes.

Objective: The aim was to evaluate the effects of a Web-based component of a weight loss service (Imperative Health) in an overweight/obese population at risk of cardiovascular disease (CVD) using a randomized controlled design and a true control group.

Methods: A total of 65 overweight/obese adults at high risk of CVD were randomly allocated to 1 of 2 groups. Group 1 (n=32) was provided with the Web-based program, which supported positive dietary and physical activity changes and assisted in managing weight. Group 2 continued with their usual self-care (n=33). Assessments were conducted face-to-face. The primary outcome was between-group change in weight at 3 months. Secondary outcomes included between-group change in anthropometric measurements, blood pressure, lipid measurements, physical activity, and energy intake at 3, 6, and 12 months. Interviews were conducted to explore participants' views of the Web-based program.

Results: Retention rates for the intervention and control groups at 3 months were 78% (25/32) vs 97% (32/33), at 6 months were 66% (21/32) vs 94% (31/33), and at 12 months were 53% (17/32) vs 88% (29/33). Intention-to-treat analysis, using baseline observation carried forward imputation method, revealed that the intervention group lost more weight relative to the control group at 3 months (mean –3.41, 95% CI –4.70 to –2.13 kg vs mean –0.52, 95% CI –1.55 to 0.52 kg, P<.001), at 6 months (mean –3.47, 95% CI –4.95 to –1.98 kg vs mean –0.81, 95% CI –2.23 to 0.61 kg, P=.02), but not at 12 months (mean –2.38, 95% CI –3.48 to –0.97 kg vs mean –1.80, 95% CI –3.15 to –0.44 kg, P=.77). More intervention group participants lost ≥5% of their baseline body weight at 3 months (34%, 11/32 vs 3%, 1/33, P<.001) and 6 months (41%, 13/32 vs 18%, 6/33, P=.047), but not at 12 months (22%, 7/32 vs 21%, 7/33, P=.95) versus control group. The intervention group showed improvements in total cholesterol,

XSL•FO
RenderX

triglycerides, and adopted more positive dietary and physical activity behaviors for up to 3 months verus control; however, these improvements were not sustained.

Conclusions: Although the intervention group had high attrition levels, this study provides evidence that this Web-based program can be used to initiate clinically relevant weight loss and lower CVD risk up to 3-6 months based on the proportion of intervention group participants losing ≥5% of their body weight versus control group. It also highlights a need for augmenting Web-based programs with further interventions, such as in-person support to enhance engagement and maintain these changes.

Trial Registration: ClinicalTrials.gov NCT01472276; http://clinicaltrials.gov/ct2/show/study/NCT01472276 (Archived by Webcite at http://www.webcitation.org/6Z9lfj8nD).

(J Med Internet Res 2015;17(7):e177) doi:10.2196/jmir.3828

KEYWORDS

Internet; randomized controlled trial; health behavior; weight loss; overweight; obesity

Introduction

The prevalence of obesity has been increasing progressively throughout the world [1]. Identifying effective and cost-effective treatment and prevention strategies is a top priority for all health care systems. Over the past few decades, the Internet has increasingly been used to deliver behavioral modification programs owing to its easy accessibility and anonymity, potential for wide reach and penetration, and its ability to provide a source of continuous support to large segments of the population [2-4].

There is growing evidence suggesting that the Internet may be a viable medium for encouraging weight loss. However, several systematic reviews and meta-analyses, conducted in this area have found it difficult to draw definitive conclusions regarding its effectiveness owing to heterogeneity in study designs, methods employed, and the lack of "true control" groups used [5-9]. Most of the evidence to date comes from randomized controlled trials (RCTs) conducted in the United States. Many only included short-term follow-up and lacked true control groups (no support provided), making it challenging to accurately evaluate the true effectiveness of Web-based programs. Instead, minimal support groups are often employed to help boost recruitment and decrease attrition; although this approach may attenuate the relationship between groups and limits the ability of the findings to inform cost-effectiveness and health care models. Research is also limited regarding the effect of these Web-based programs on other health outcomes that coexist with weight loss, such as cardiovascular disease (CVD) risk factors. Therefore, the aim of this study was to evaluate the effects of an interactive Web-based component of a service called Imperative Health on weight loss (primary outcome) and CVD risk factors (secondary outcomes) in an overweight and obese population at high risk of CVD using a randomized controlled design and a true control group. It was hypothesized that weight loss would be greater in the Web-based program intervention group compared to the usual care control group.

Methods

Recruitment

Ethical approval was obtained from the Office for Research Ethics Committees Northern Ireland. The trial was registered (ClinicalTrials.gov identifier: NCT01472276) and is reported in accordance with the Consolidated Standards of Reporting Trials (CONSORT)-eHealth checklist (see Multimedia Appendix 1) [10]. Participants were recruited from April to December 2011 using posters in public places in the greater Belfast area and intranet advertisements via staff updates in the Belfast Health and Social Care Trust and Queen's University Belfast (QUB). Patients from the Regional Centre for Diabetes and Endocrinology at the Royal Victoria Hospital Belfast were also sent a letter informing them about the study. Participants were eligible if they were older than 18 years, had a body mass index (BMI) between 27 and 40 kg/m^2, were inactive or moderately inactive assessed by the General Practice Physical Activity Questionnaire (GPPAQ) [11] and had 1 or more CVD risk factors: high blood pressure ≥140/90 mmHg, total cholesterol ≥5.0 mmol/L, or type 2 diabetes mellitus. All participants were required to have access to the Internet, email, and a telephone and were asked not to participate in another behavioral change weight loss program throughout the study period. Participants were excluded if they had established CVD, type 1 diabetes mellitus, were pregnant, or consumed excessive amounts of alcohol. Computer literacy was not assessed. All participants at the screening appointment provided written informed consent.

Study Design

After completion of the baseline assessments, conducted face-to-face at the Regional Centre for Diabetes and Endocrinology at the Royal Victoria Hospital Belfast, participants were randomly allocated to 1 of 2 parallel groups (1:1 allocation ratio) using a block randomization approach (block size=10) with computer-generated numbers. A researcher independent from the study prepared the randomization schedule. Opaque sealed envelopes were used to conceal the sequence until groups were allocated. Participants were recruited and enrolled by the researcher, who was unaware of the randomization schedule until after the baseline assessments when the sealed envelope containing the allocation outcome was opened by the participant. Group 1 (intervention group) was provided with the Web-based program, known as Imperative Health, excluding telephone and email support and group 2 (control group) were requested to continue with their usual self and medical care. All participants were followed up at 3, 6, and 12 months after randomization for assessment of primary and secondary outcomes. Based on a standard deviation of weight loss at 3 months of 3.0 kg observed in a number of

XSL•FO
RenderX

Internet-based weight loss studies in the literature [12-15], it was estimated that a sample size of 60 (30 per group) would give the study 90% power at the 5% significance level to detect a difference of 2.6 kg between groups at the 3 month follow-up. Allowing for a 10% dropout rate at the first 3-month follow-up, we aimed to recruit 66 participants. With only one researcher on the ground, it was not possible to blind the researcher or participants to group allocation, but laboratory analysis was performed blind.

Intervention (Imperative Health Web-Based Program)

Overview

Imperative Health is a service owned by AXA PPP Healthcare Limited that consists of a Web-based program and human (email and telephone) support that assists in lifestyle change, with a particular focus on improving diet and nutrition, increasing physical activity, and managing weight and other CVD risk factors. It combines objective monitoring of weight and physical activity with automated, tailored feedback and support by physiologists by telephone and email. Previous versions of this Web-based program have been evaluated by Hurling et al [16,17] and Ware et al [18]. This program has since been modified to be more relevant to individuals with independent risk factors for CVD, such as hypertension, dyslipidemia (high cholesterol and triglycerides), and type 2 diabetes mellitus. For this particular study, only the Web-based program component of the service was evaluated to determine its specific impact (ie, the human support [telephone and email] component of the service was removed for the purposes of this trial).

Initial Setup of Imperative Health (Web-Based Program)

At the end of the baseline appointment, the intervention group participants were provided with the Imperative Health package that contained the self-monitoring devices (Bluetooth-enabled weighing scales and an accelerometer activity band) and basic written instructions to set up an online account at home. To access the online program, participants were instructed to go to the Imperative Health website [19] and enter a unique code to create their own personal password-protected free account. Participants were advised to follow the online instructions to complete registration and to enable the setup of the monitoring devices. The intervention group was informed at the baseline appointment that if any problems regarding the technology occurred throughout the study period after the initial setup, they were to contact Imperative Health rather than the researcher. This study wanted to evaluate this Web-based program in a real-life setting to determine realistic levels of engagement and their relationship with weight loss; therefore, no instructions were provided by the researcher as to how often the participants should log in to use the website components and the self-monitoring devices. The Web-based program, however, does encourage daily engagement by allowing the upload of daily weight and physical activity data and by the entry of daily food diaries (described in detail subsequently).

Web-Based Behavior Change Program

Once the online account was set up, the participants were required to complete a series of online introductory health questionnaires that enabled Imperative Health to collect information on their height, weight, waist circumference, blood pressure and blood biomarkers (total cholesterol, high-density lipoprotein [HDL] cholesterol, fasting blood glucose, and triglycerides), as well as information on past and current health status, dietary intake, physical activity level, and stated goals. This self-reported information was not used by the researcher to evaluate the effects of this Web-based program; instead, it was used by the Imperative Health system to generate personalized daily targets (weight loss, physical activity, and dietary targets) for each participant to achieve over 12 weeks. Automated weekly feedback on their performance, assessed by the self-monitoring devices (weighing scales and accelerometer) and the food diary was provided, and also in the form of an overall review after 12 weeks. After 12 weeks, to encourage further progress, it was requested that the participants start a new program by completing the same introductory health questionnaires again and setting new goals. The Web-based program encompassed supportive components to help facilitate lifestyle change (see Table 1). These components of the Web-based program were developed based on well-recognized behavior change strategies, such as planning, self-monitoring, goal setting, and structured feedback, which were all used within the Diabetes Prevention Program [20] to promote weight loss.

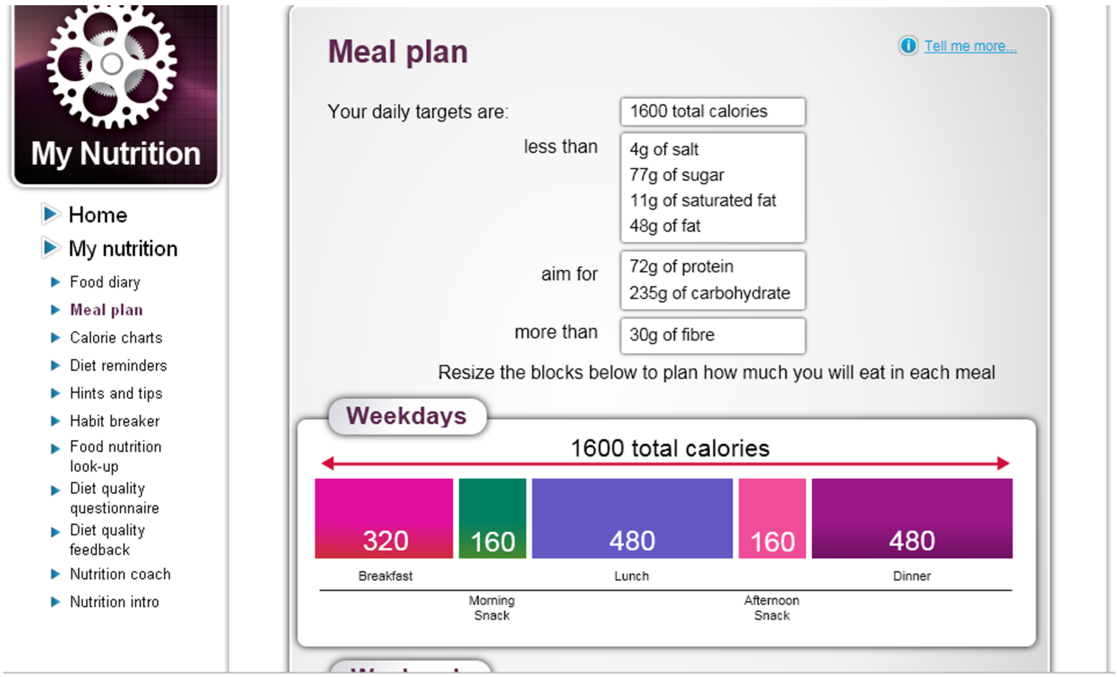
XSL•FO
RenderX

Table 1. Imperative Health Web-based program components to support behavior change.

Behavior change strategy and Web-based program component	Description of component
Goal setting	
Daily dietary targets, daily physical activity targets, weekly weight loss targets, clinical targets	Personalized daily dietary, physical activity (for screenshots see Figure 1), weight, and clinical (blood pressure, glucose, lipids) targets were created based on the health questionnaire responses. Targets were reviewed every 12 weeks.
Planning	
Exercise weekly schedule	A weekly schedule for planning physical activity was provided. Icons (representing light, moderate, or vigorous activities) could be dragged to specific days. Start times and duration of the activity could be selected (for screenshots see Figure 1).
Daily meal planners	Meal suggestions for breakfast, lunch, dinner, and snacks were provided to help meet personalized dietary targets set by the Web-based program.
Self- monitoring	
Bluetooth weighing scales, Bluetooth accelerometer activity band	Monitoring devices included Bluetooth-enabled weighing scales and an activity band. Data from the weighing scales was transmitted to the activity band and subsequently sent to the user's online profile page. The activity band provided daily feedback on minutes of moderate, high, and very high activity (for screenshots see Figure 2).
Food diary (calorie uploads), clinical measurements (blood pressure, glucose, blood lipids uploads)	Daily calorie intake, blood pressure, glucose, and blood lipid measurements could be entered and uploaded onto colored charts (for screenshots see Figures 2 and 3) to demonstrate daily, weekly, and monthly results and if targets were achieved.
Personalized feedback	
Coaching session, automated weekly feedback	Automated tailored feedback on progress was provided weekly.
Push reminders	
Email/SMS texts	Text messages or emails were sent daily and weekly to help remind participants to log in and to weigh themselves.
Social support	
Community forum	Online discussion forums were available for comments to be posted.
Decisional balance theory	
Habit breaker component	Solutions for barriers perceived as preventing healthier behaviors (eg, eating breakfast) being adopted were provided.

XSL•FO
RenderX

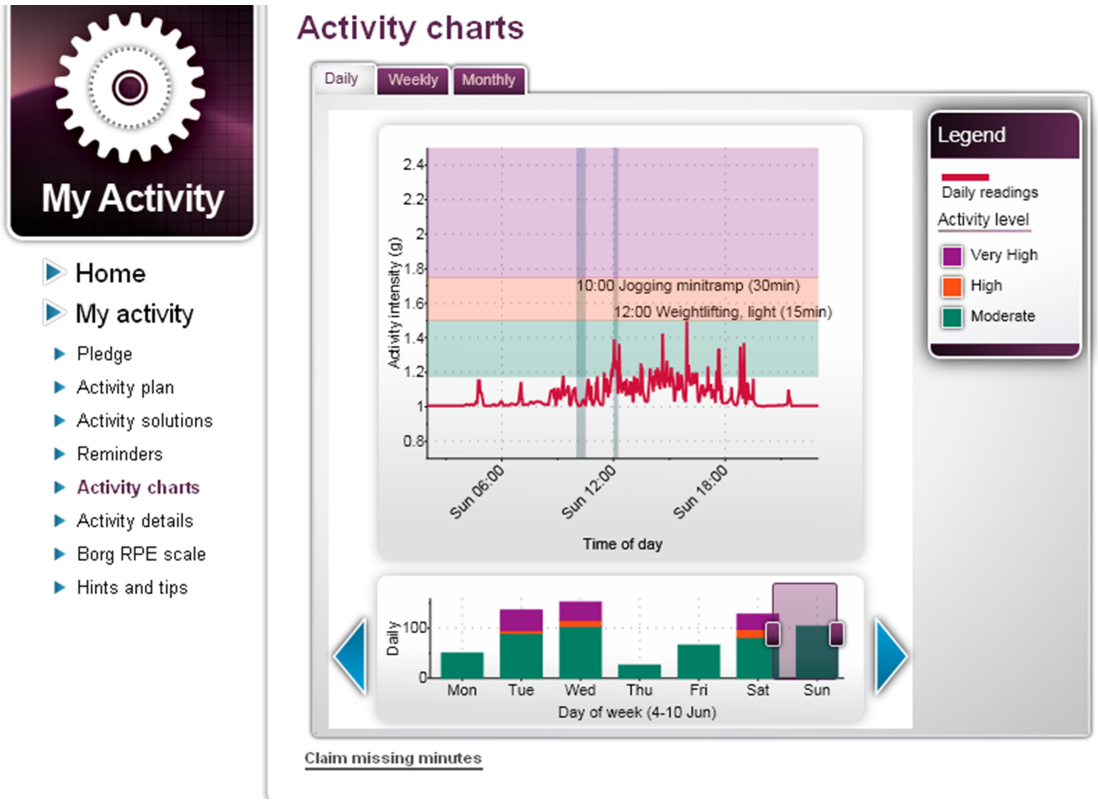
Figure 1. Imperative Health screenshots of meal and activity planners.

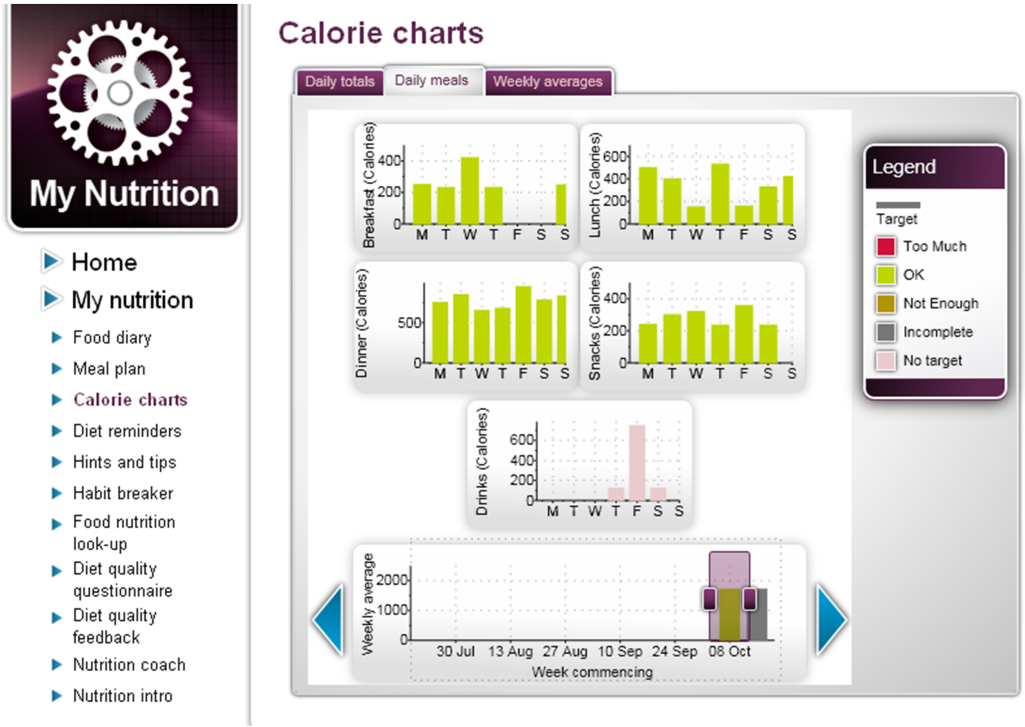




XSL•FO
RenderX

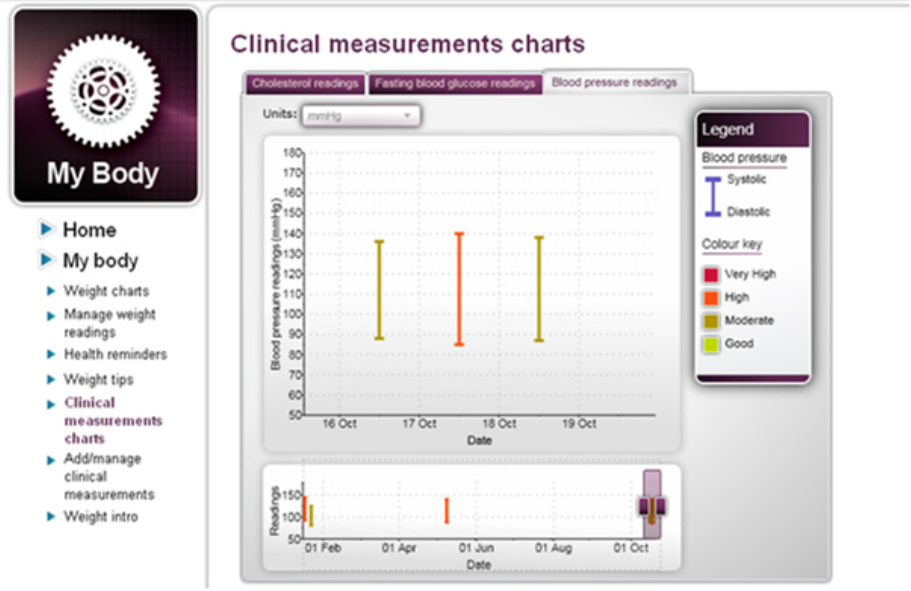
Figure 2. Imperative Health screenshots of activity and calorie feedback charts.





XSL•FO
RenderX

Figure 3. Imperative Health screenshot of clinical measurement feedback chart.



Outcome Measures

Overview

Assessments were carried out face-to-face at the Regional Centre for Endocrinology and Diabetes at the Royal Victoria Hospital Belfast at baseline, 3 months, 6 months, and 12 months.

Primary Outcome

The primary outcome for this study was between-group change in body weight (kg) at 3 months. Weight was measured, without shoes and in light clothing, to the nearest 0.1 kg using calibrated Salter 994 digital weighing scales (Salter Housewares Ltd, Tonbridge, UK).

Secondary Outcomes

Secondary outcomes were between-group change in weight loss at 6 and 12 months, and between-group change in the following risk markers at each follow-up: BMI calculated as weight (kg) divided by height squared (m^2); height was measured to the nearest 0.1 cm using a Leicester portable height measure (CMS Weighing Equipment Ltd, London, UK); waist circumference was measured to the nearest 0.5 cm using a tape measure at the middle point between the lower rib margin and iliac crest at normal expiration.

Blood pressure (mm Hg) was measured using an automated Omron M3 sphygmomanometer (Omron Healthcare, Hoofddorp, The Netherlands).

Fasting serum lipid profile included measurements of total cholesterol, HDL cholesterol, and triglycerides and were measured using standard assays on an automated ILab 600 Chemistry system (Instrumentation Laboratory, Cheshire, UK). Plasma high-sensitivity C-reactive protein (hs-CRP) was measured using an ultrasensitive assay (Quantex CRP Ultrasensitive; Instrumentation Laboratory, Cheshire, UK) on an automated machine (ILab 600 Chemistry System).

Dietary intake was assessed using a diet history interview [21], which was a retrospective dietary assessment method used to gather information regarding the habitual food intake of all participants over the previous 3 months. The diet history method has been shown to have good repeatability in previous studies and is also able to pick up dietary changes over time [22]. Quantities of food and food portion sizes (household measures) were converted into weights (grams) by using Crawley's Food Portion Sizes (Food Standards Agency) [23]. The food type, preparation method if relevant, and weight of food were entered into a computerized food analysis database (WISP, Weighed Intake Software Program; Tinuviel Software, Warrington, UK). For the purpose of this study, total daily energy intake (kcal) was calculated.

Physical activity was assessed using the validated Recent Physical Activity Questionnaire (RPAQ) [24]. Participants were asked to provide descriptions of their habitual physical activity performed in 4 domains: home, work, travel, and recreation over the last 4 weeks. For the purpose of this study, time (min/day) spent participating in moderate and vigorous activities (> 3.5 Metabolic Equivalent Task, MET) was calculated.

A self-reported questionnaire was distributed at the baseline appointment to collect sociodemographic information including past and current occupation. Socioeconomic status was classified according to National Statistics Socio-economic Classification (NE-SEC) [25] into 3 occupational classes: the highest included higher managerial, administrative, and professional occupations; the second class was intermediate occupations; and the third class included routine and manual occupations.

Website Usage

Data on frequency of log-ins, the total number of completed food diaries, and the number of weight and physical activity uploads from the monitoring devices were provided by Imperative Health and were used to determine level of engagement.

Qualitative: Interviews

To gain in-depth feedback on the intervention group's experiences of using the Web-based program, these participants

XSL•FO
RenderX

were asked if they would be willing to take part in an interview conducted by the researcher toward the end of the study. This was an optional part of the study; therefore, a convenience sampling technique was utilized. The interviews were conducted between July and August 2012, in the Centre for Public Health, QUB, within an informal setting and lasted approximately 25 to 30 minutes. Semistructured open-ended questions were used throughout to ensure that a consistent approach was utilized. The researcher used a style of probing to extract more information or clarify meaning.

All the interviews were audio-recorded and transcribed verbatim. NVivo 8 was used to assist in the management and analysis of the transcripts. To analyze the transcripts, a template approach, outlined by Crabtree and Miller [26], was utilized. This process involved the naming, defining, and describing of the codes based on research questions. Three broad categories formed the code template: views on their experiences of using Imperative Health (Web-based program), views on Imperative Health's website components that support behavior change, and suggested improvements that Imperative Health should implement. The template of codes was then applied to all transcripts. Given that the data were qualitative, frequencies were used in the broadest sense (eg, majority, some, and few). Quotations were used to demonstrate typical views within each code category.

Statistical Analysis

All analyses were performed using SPSS for Windows version 21.0 (SPSS Inc, Chicago, IL, USA). Results are expressed as mean and standard deviation for normally distributed variables and median and interquartile range for variables that did not satisfy normality criteria. Categorical data are expressed as frequencies and percentages. To compare baseline characteristics between the control and intervention groups, for continuous variables, the appropriate parametric (independent samples *t* test) and nonparametric tests (Mann-Whitney *U* test) were utilized. For categorical variables, the chi-square test was used. Between-group differences in the primary outcome (weight change at 3 months) and secondary outcomes from baseline to 3 months, 6 months, and 12 months were investigated using the analysis of covariance (ANCOVA) adjusted for baseline measurements [27]. Analyses were carried out by an intention-to-treat (ITT) approach using a single imputation method (baseline observation carried forward, BOCF) to deal with missing data and losses to follow-up [28]. A complete-case analysis on weight change was also conducted using information on all individuals with available data at each time point. A sensitivity analysis was conducted using MANOVA, but this reached similar findings (results not shown). Triglycerides, CRP, and physical activity distributions were skewed; therefore, they were log transformed. Adjusted differences in log-transformed means between groups from ANCOVA were converted to, and reported as, ratios of geometric means and 95% confidence intervals. Within-group changes (intervention or control) in weight loss were analyzed using paired-sample *t* tests. Because the Web-based program usage data was not normally distributed, Spearman correlations were performed to investigate the relationship between weight change and Web-based program usage at each time point (intervention group only).

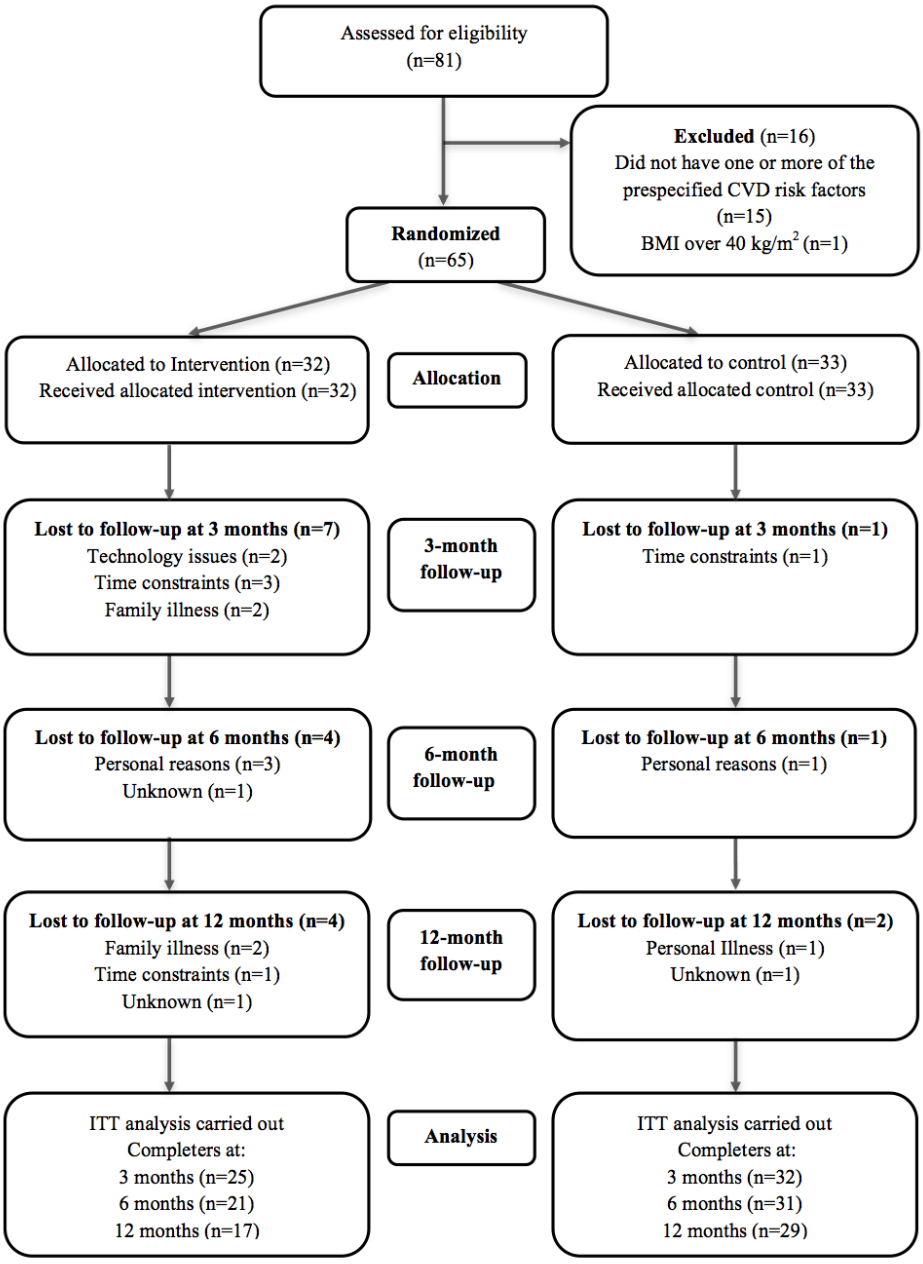
Results

Participant Flow

A total of 81 individuals were screened for eligibility; 16 were ineligible and the other 65 participants (29 males, 36 females) were randomized to the control (n=33) or intervention (n=32) groups (see Figure 4). Retention rates significantly differed between the control and intervention groups at 3 months (32/33, 97% vs 25/32, 78%, *P*=.03), at 6 months (31/33, 94% vs 21/32, 66%, *P*=.004), and at 12 months (29/33, 88% vs 17/32, 53%, *P*=.002), respectively. Baseline characteristics, including socioeconomic status, did not differ significantly between those who dropped out of the study and those who completed the study in either group at any time point.

XSL•FO
RenderX

Figure 4. CONSORT diagram showing the flow of participants through the trial and analyzed for weight loss at 3 months, 6 months, and 12 months.



Baseline Characteristics

Mean age was 52.1 (SD 7.4) years and mean BMI at baseline was 32.7 (SD 2.9) kg/m^2. According to World Health Organization criteria [29], 20% (13/65) of the sample were overweight, 58% (38/65) were obese category I, and 22% (14/65) were obese category II. Socioeconomic status (SES) was determined by occupational class (NS-SEC): Class 1 included higher managerial, administrative, and professional occupations, which 51% (33/65) of the sample lay within; Class 2 included intermediate occupations of which 38% (25/65) of the sample lay within; and Class 3 included routine and manual occupations, and applied to 11% (7/65) of the sample. There were no significant differences in any of the baseline characteristics between the control and intervention group (see Table 2).

XSL•FO
RenderX

Table 2. Baseline characteristics of intervention and control participants according to control and intervention group.

Characteristic[a]	Intervention (n=32)	Control (n=33)	*P* [b]
Demographics			
Gender, n (%)			.39
Male	16 (50)	13 (39)	
Female	16 (50)	20 (61)	
Age (years), mean (SD)	51.4 (7.59)	52.9 (7.27)	.43
Physical measurements, mean (SD)			
Weight (kg)	95.2 (16.7)	91.9 (13.4)	.39
Height (cm)	169.4 (9.44)	168.1 (9.35)	.57
BMI (kg/m^2)	32.9 (3.07)	32.4 (2.74)	.50
Waist circumference (cm)	103.5 (11.2)	102.5 (9.47)	.69
Blood pressure (mm Hg), mean (SD)			
Systolic	129.8 (17.8)	129.1 (18.3)	.88
Diastolic	85.5 (9.54)	86.0 (11.4)	.85
Blood biomarkers			
Total cholesterol (mmol/L), mean (SD)	4.87 (1.44)	5.16 (1.02)	.35
HDL (mmol/L), mean (SD)	1.33 (0.39)	1.36 (0.31)	.77
Triglycerides (mmol/L), median (IQR)	1.49 (1.18-1.86)	1.48 (1.01-2.02)	.75
CRP (mg/L), median (IQR)	1.73 (0.67-2.90)	2.11 (1.11-4.35)	.22
Energy intake and physical activity			
Energy (kcal), mean (SD)	1949.6 (545.1)	1893.6 (477.2)	.66
Physical activity (min/day), median (IQR)	15.5 (6.4-45.3)	17.4 (7.5-46.9)	.72

[a] BMI: body mass index; HDL: high-density lipoprotein cholesterol; CRP: C-reactive protein.

[b] Between-group differences analyzed using independent samples *t* test for normal data and Mann-Whitney U test for skewed data. Differences between categories analyzed using chi-square test.

Change in Body Weight (Primary Outcome)

As shown in Table 3, both approaches (ITT and complete case) demonstrated significant mean weight loss difference between groups at 3 months; however, the magnitude of weight lost was slightly higher using the complete-case analysis approach. ITT analysis revealed that the intervention group participants had a mean weight loss of –3.41 kg at 3 months; the control group lost –0.52 kg. Overall, this accounted for a significant mean weight difference between groups of –2.70 kg after adjusting for baseline weight (P<.001).

XSL•FO
RenderX

Table 3. Weight (kg) outcome differences between and within study groups from baseline to 3, 6, and 12 months (intention-to-treat [ITT] and complete-case analysis).

Analysis and month	Change from baseline, mean (95% CI)[a]		Difference between groups[b]	
	Intervention	Control	Mean (95% CI)	*P*
ITT [c]				
3	–3.41 (–4.70, –2.13)***	–0.52 (–1.55, 0.52)	–2.70 (–4.27, –1.13)	.001
6	–3.47 (–4.95, –1.98)***	–0.81 (–2.23, 0.61)	–2.49 (–4.50, –0.48)	.02
12	–2.38 (–3.48, –0.97)**	–1.80 (–3.15, –0.44)*	–0.27 (–2.16, 1.61)	.77
Complete case [d]				
3	–4.37 (–5.80, –2.94)***	–0.53 (–1.60, 0.54)	–3.66 (–5.28, –2.05)	<.001
6	–5.28 (–7.12, –3.44)***	–0.86 (–2.38, 0.65)	–4.16 (–6.46, –1.86)	.001
12	–4.48 (–7.34, –2.37)**	–2.16 (–4.58, –0.62)*	–1.89 (–4.42, 0.64)	.14

[a] Within-group weight changes were analyzed using paired-sample *t* tests and only significant results are presented. *$P<.05$, **$P<.01$, ***$P<.001$.

[b] Difference between groups analyzed using ANCOVA and adjusted for baseline weight.

[c] ITT analysis: control group (n=33) and intervention group (n=32) at 3, 6, and 12 months.

[d] For the complete-case analysis: control group (n=32) at 3 months, (n=30) at 6 months, and (n=29) at 12 months. Intervention group (n=25) at 3 months, (n=21) at 6 months, and (n=17) at 12 months.

Change in Body Weight at 6 and 12 Months (Secondary Outcome)

The ITT analysis (see Table 3) demonstrated that the intervention group lost significantly more weight compared to the control group from baseline to 6 months (mean –3.47, 95% CI –4.95 to –1.98 kg vs mean –0.81, 95% CI –2.23 to 0.61 kg; *P*=.02, respectively), but not from baseline to 12 months (mean –2.38, 95% CI –3.48 to –0.97 kg vs mean –1.80, 95% CI –3.15 to –0.44 kg; *P*=.77). There were significant changes in weight between baseline and each time point within the intervention group (3 months: *P*<.001; 6 months: *P*<.001; 12 months: *P*=.002). However, between 6 months and 12 months the intervention group gained 1.08 kg, reducing the overall mean weight loss at 12 months in this group. There was a significant weight loss from baseline to 12 months within the control group (mean –1.80, 95% CI –3.15 to –0.44 kg, *P*=.01), but not for the 3 month (*P*=.32) and 6 month (*P*=.25) time points.

Percentage Weight Loss

Weight loss as a percentage of baseline weight was calculated using the ITT data. The mean percentage weight loss in the intervention and the control group was from baseline to 3 months mean –3.62% (95% CI –4.95 to –2.29) vs mean –0.34% (95% CI –1.34 to 0.65), respectively (*P*<.001); from baseline to 6 months mean –3.73% (95% CI –5.30 to –2.16) vs mean –0.63% (95% CI –2.06 to 0.80), respectively (*P*=.004); and from baseline to 12 months mean –2.42% (95% CI –3.93 to –0.91) vs –1.94% (95% CI –3.26 to –0.39), respectively (*P*=.56). Significantly more participants in the intervention group compared to the control group lost 5% or more of their baseline body weight at 3 months (11/32, 34% vs 1/33, 3%, *P*<.001) and at 6 months (13/32, 41% vs 6/33, 18%, *P*=.047), but not at 12 months (7/32, 22% vs 7/33, 21%, *P*=.95).

Change in Other Secondary Outcomes

Table 4 shows the intervention group significantly reduced their BMI and waist circumference measurements relative to the control group from baseline to 3 months (*P*<.001 and *P*=.006, respectively) and to 6 months (*P*=.003 and *P*=.02, respectively), but not at 12 months. There were no between-group differences in blood pressure observed during the study. For lipid measurements, larger reductions were observed in total cholesterol and triglyceride concentrations in the intervention group compared to the control group, but only during the first 3 months (*P*=.003 and *P*=.003, respectively). Similar patterns were identified for health behaviors: the intervention group significantly decreased their energy intake and increased their time spent exercising at an intensity greater than 3.5 METs relative to the control group from baseline to 3 months (*P*=.005 and *P*=.03). These behaviors were not sustained over the longer term at 6 and 12 months.

XSL•FO
RenderX

Table 4. Clinical outcome differences between study groups from baseline to 3, 6, and 12 months (intention to treat).

Clinical outcome and month[a]	Change from baseline, mean (95% CI)		Between-group difference	
	Intervention (n=32)	Control (n=33)	Adjusted mean (95% CI)[b]	*P*
BMI (kg/m^2)				
3	–1.16 (–1.60, –0.73)	–0.14 (–0.47, 0.19)	–0.99 (–1.53, –0.46)	<.001
6	–1.20 (–1.70, –0.70)	–0.18 (–0.64, 0.27)	–1.02 (–1.69, –0.35)	.003
12	–0.78 (–1.26, –0.31)	–0.65 (–1.12, 0.19)	–0.10 (–0.75, 0.55)	.76
Waist circumference (cm)				
3	–2.73 (–3.98, –1.49)	–0.67 (–1.44, 0.11)	–2.04 (–3.47, –0.61)	.006
6	–3.05 (–4.68, –1.41)	–0.83 (–1.95, 0.28)	–2.18 (–4.11, –0.24)	.02
12	–2.31 (–3.84, –0.79)	–1.80 (–3.02, –0.58)	–0.42 (–2.29, 1.45)	.66
Systolic blood pressure (mm Hg)				
3	–2.69 (–6.48, 1.10)	–1.64 (–6.02, 2.75)	–0.81 (–5.61, 3.99)	.74
6	–1.31 (–4.83, 2.20)	0.88 (–3.79, 5.55)	–1.92 (–6.48, 2.65)	.40
12	–1.22 (–4.33, 1.90)	–2.12 (–2.25, 6.49)	–3.13 (–7.69, 1.43)	.18
Diastolic blood pressure (mm Hg)				
3	–3.03 (–5.14, –0.92)	–2.36 (–5.02, 0.29)	–0.83 (–3.76, 2.10)	.58
6	–2.63 (–5.05, –0.20)	–1.73 (–5.26, 1.81)	–1.14 (–4.55, 2.27)	.51
12	–1.78 (–3.52, –0.05)	–1.55 (–4.57, 1.48)	–0.38 (–3.52, 2.76)	.81
Total cholesterol (mmol/L)				
3	–0.49 (–0.70, –0.28)	–0.06 (–0.31, 0.19)	–0.48 (–0.79, –0.18)	.003
6	–0.30 (–0.53, –0.08)	–0.24 (–0.46, –0.02)	–0.07 (–0.38, 0.24)	.64
12	–0.19 (–0.38, –0.01)	–0.13 (–0.36, 0.10)	–0.09 (–0.38, 0.20)	.56
HDL (mmol/L)				
3	–0.02 (–0.08, 0.04)	0.00 (–0.07, 0.07)	–0.03 (–0.11, 0.06)	.51
6	–0.01 (–0.07, 0.06)	–0.03 (–0.10, 0.04)	0.02 (–0.07, 0.12)	.62
12	–0.02 (–0.07, 0.02)	0.02 (–0.06, 0.10)	–0.04 (–0.13, 0.04)	.32
Triglycerides (mmol/L) [c]				
3	0.89 (0.82, 0.96)	1.03 (0.96, 1.10)	0.87 (0.80, 0.95)	.003
6	0.96 (0.89, 1.04)	0.98 (0.91, 1.04)	0.99 (0.90, 1.09)	.79
12	0.97 (0.91, 1.02)	0.97 (0.90, 1.05)	1.00 (0.92, 1.09)	.93
CRP (mg/L) [c]				
3	0.93 (0.84, 1.03)	1.01 (0.87, 1.18)	0.88 (0.74, 0.96)	.13
6	0.89 (0.83, 0.95)	1.03 (0.86, 1.24)	0.83 (0.69, 1.01)	.06
12	0.99 (0.80, 1.23)	0.99 (0.81, 1.21)	0.93 (0.71, 1.21)	.58
Energy intake (Kcal)				
3	–487.6 (–640.7, –334.5)	–241.4 (–375.8, –106.9)	–216.3 (–364.0, –68.7)	.005
6	–314.8 (–466.2, –163.5)	–243.2 (–393.6, –92.8)	–47.8 (–228.5, 133.0)	.60
12	–221.6 (–363.5, –79.8)	–204.2 (–384.5, 23.9)	7.20 (–191.0, 205.4)	.94
Physical activity (min/day) [c,d]				
3	2.85 (1.64, 4.94)	1.43 (0.92, 2.21)	1.98 (1.09, 3.60)	.03
6	1.43 (1.00, 2.06)	1.00 (0.61, 1.64)	1.43 (0.84, 2.44)	.19
12	1.52 (0.95, 2.43)	1.18 (0.74, 1.89)	1.28 (0.72, 2.27)	.40

XSL•FO
RenderX

[a] BMI: body mass index; HDL: high-density lipoprotein cholesterol; CRP: C-reactive protein.

[b] Difference between groups analyzed using ANCOVA and adjusted for baseline measurements.

[c] Data presented as ratio of geometric mean and 95% confidence for log-transformed variables.

[d] Physical activity calculated as time in minutes spent exercising >3.5 METs daily.

Website Usage and Weight Change (Intervention Group Only)

Website utilization data are presented in Table 5. Participants in the intervention group tended to log in, upload their weight measurement, and make food diary entries more frequently during the first 3 months of the intervention; website usage declined thereafter.

Table 5. Website utilization patterns (intervention group only) from baseline to 3, 6, and 12 months.

Website components	Baseline to 3 months (13 weeks), median (IQR)[a]	3 to 6 months (13 weeks), median (IQR)[a]	6 to 12 months (26 weeks), median (IQR)[a]
Number of log-ins	69.0 (25.5-122.0)	12.0 (2.0-47.5)	27.0 (2.0-96.8)
Food diary entries	18.0 (0.0-77.0)	0.0 (0.0-38.5)	0.0 (0.0-156.5)
Weight uploads	15.0 (9.0-46.0)	11.0 (1.0-30.0)	4.0 (0.0-29.8)
Physical activity uploads	13.0 (10.0-13.0)	12.0 (3.0-13.0)	15.0 (0.0-24.5)

[a] Data presented as median (IQR) due to data being skewed. Sample size at 3 months (n=25), at 6 months (n=21), and at 12 months (n=17).

Correlation analyses (see Table 6) demonstrated that weight change from baseline to 3-month follow-up was significantly positively related to the number of log-ins (*P*=.04) and the number of weight uploads (*P*=.007) at 3 months. A positive relationship was observed between weight change from baseline to 6 months and the amount of physical activity uploads over the same time period (*P*=.048). The number of daily food diaries entered was not related to weight change throughout the course of the study.

Table 6. Spearman's correlations (ρ) between weight change and website components usage from baseline (intervention group only).

Website component usage	Weight change from baseline					
	3 months		6 months		12 months	
	ρ	*P*	ρ	*P*	ρ	*P*
Number of log-ins	.42	.04	.28	.21	.21	.42
Food diary entries	.01	.96	.00	.99	–.20	.53
Physical activity uploads	.33	.14	.47	.048	.12	.67
Weight uploads	.53	.007	.20	.39	.10	.70

Interview Feedback (Intervention Group Only)

Overview

A total of 7 participants (4 males and 3 females) from the intervention group were recruited using a convenience sampling approach. Three broad categories formed the code template: views on their experiences of using Imperative Health (Web-based program), views on Imperative Health's website components (see Table 1) that support behavior change, and suggested improvements that Imperative Health should implement. Presented subsequently are some of the quotations used to demonstrate typical views within each code category.

Experiences Using Imperative Health (Web-Based Program)

All interviewees stated that they were keen to use this Web-based program to help them lose weight and manage their chronic condition. They found the initial setup of their Imperative Health accounts relatively straightforward:

> *It was very straightforward. I don't think I had any difficulty at all with it.*

Some of the interviewees perceived using the Web-based program as time consuming and quite burdensome; specifically, the tasks that involved uploading and manually entering measurements as well as working through the weekly feedback:

> *I don't know whether people who would be employed full time would have enough time to do that...If you're running out to work in the morning and you have to be out for 7 o'clock, you're not going to be standing there weighing yourself, and again, when you come back home again, typing in what you have done.*

> *If you wanted a quick consult, it was taking you 10 minutes to get through...*

Website Components That Support Behavior Change

The majority of interviewees found the personalized targets (weight loss, physical activity, and dietary targets) provided by the Web-based program realistic and motivating:

XSL•FO
RenderX

> *It did give me the motivation to say, "Right, I'm supposed to go walking 60 minutes a day. I'll try and keep to that target of 60 minutes a day." They weren't tough; the calories I was being allowed were okay...*

The majority of interviewees were not impressed with the Web-based program planning components (meal planner and exercise schedule). Accessibility issues and aversions to the recommended foods in the meal plans were commented upon:

> *If it had of been prepackaged food or something that I would have actually liked but there was none of the stuff that really appealed to me. So I never used the meal planner...*
>
> *But they were all foods that were for supermarkets, say, in England and a lot of the stuff that you wouldn't get here maybe.*

All the interviewees felt that the Web-based program's self-monitoring components (weighing scales and accelerometer activity band—data from these devices were uploaded onto colored charts to track progress) helped them to evaluate their progress and at the same time acted as facilitators for motivating them to keep continuing toward their targets.

> *With regard to the weight one, it encouraged you to do better, because it showed if you were flattening out or, at worst, going the wrong way off your target.*
>
> *I found the activity useful because when I would sync up at the end of the week I would have a look and say "I was low on Tuesday and Wednesday this week. I'll maybe do a boost on Friday. I'll go for an hour and a half walk just to make sure my average for the week is up." So I found that a little bit, slightly motivating.*

All interviewees provided negative feedback regarding the dietary self-monitoring element of this Web-based program (food diary). They felt the process was time consuming and burdensome as a result of having to look up all the calories of the foods they consumed and then enter them manually into the food diary:

> *I found getting the nutritional values of things awkward because you had to go into a separate wee thing in the background and then you had to write it down and then you go back to something else.*

Most of the interviewees claimed they did not use the component for monitoring their clinical measurements, as they were unable to get these health risk factors measured regularly:

> *...The average person doesn't have that information. I might get that done twice a year.*

The majority of the participants were not impressed with the automated feedback and coaching sessions provided weekly. They felt that it was too generic and repetitive; hence, not encouraging or constructive:

> *The other thing that irritated me intensely is the standard messages that you would get at every stage of the bloody feedback! I suppose it's a computer system, what can you expect, but I just got cheesed off because it said the same thing all the time.*
>
> *It was more generic in the sense. They were just basically saying "we haven't got enough information" or "you have not met your target."*

Most participants stated that they did browse the community forum but did not contribute anything. They generally felt that there was not enough activity:

> *I occasionally dipped in and out to see what it was but there was very little action or interest, and I don't get involved in anything like that at all.*

Suggested Improvements

A common suggestion for improvement was more personalized interaction and feedback specifically from a human rather than an automated machine because this may provide them with more focus and motivation. This would be in-line with the full Imperative Health service that has physiologists supporting participants by telephone and email.

> *...some more personalized interaction in terms of somebody perhaps phoning you on your mobile to give you a kick-start or perhaps an email...*
>
> *I tried at the start, and because there is not actually a person involved in it you're not worried about what the machine tells you then, you don't care what it says to you. So you go off track a wee bit...*

Discussion

Weight Loss (Primary and Secondary Outcome)

In comparison to a "true" control group, access to a Web-based program resulted in significantly greater weight loss in the intervention group after 3 and 6 months. However, longer-term follow-up indicated that the difference in weight loss between the intervention and control group was not sustained at 12 months. The reasons for this were twofold: weight regain in the intervention group between the 6- and 12-month time point and an increase in weight loss in the control group over the same time period. In terms of clinically significant weight loss (weight loss of ≥5% of baseline body weight), significantly more participants in the intervention group compared to the control group lost 5% or more of their baseline body weight at 3 months (34% vs 3%, $P<.001$) and at 6 months (41% vs 18%, $P=.047$), but not at 12 months (22% vs 21%, $P=.95$).

Engagement, Nonusage Attrition, and Attrition

This study was designed to evaluate this Web-based program in a real-life setting to observe real levels of engagement and their relationship with weight loss; hence, no instructions were given to participants regarding how often they should log in to use the website and the self-monitoring devices. The Imperative Health program does encourage daily engagement by allowing the upload of daily weight and physical activity data, captured by the accelerometer activity band and by the entry of daily food diaries. Some studies have suggested that an unstructured self-care approach may limit the potential benefit of Internet programs [30,31]; however, prescriptive dosage studies are likely to represent efficacy rather than effectiveness and do not help to understand the likely true public health impact of these novel modes of delivery. Studies that have provided dosage

XSL•FO
RenderX

instructions have found positive effects. For example, participants that comply with the dosage instructions tend to lose significantly more weight than noncompliers [14,32-34]. The majority of these prescriptive dose studies, however, were conducted over the short term (6 months or less). Sustaining engagement levels in the long term is undoubtedly more of a challenge. Weight change (3 months) in this study had a positive moderate correlation with the number of log-ins and weight uploads, but engagement levels tended to diminish with time, particularly after 6 months. Web-based programs in general tend to have problems with long-term sustainability and nonusage attrition tends to be a common characteristic that increases steadily over time [2,31]. Participants are likely to disengage over time, perhaps due to motivational issues and, particularly, if they are failing to lose weight or have reached a plateau [35]. Furthermore, depending on the Web-based program itself and what it has to offer in terms of interactivity and level of intensity, participants may simply get bored and lose interest in the Web-based program.

Attrition rates are generally high in Web-based weight loss studies and have been reported to range between 0% and 70%, with a mean attrition rate of 22.5% [7]. Furthermore, attrition rates have been reported to be higher within the Web-based intervention group [30,32,36-38] relative to the control group, as was the case in this study.

Interactivity is essential for high engagement and low attrition; furthermore, it is well-accepted that Web-based programs with enhanced interactive features promote greater weight reduction than those that provide information only [5]. The Web-based program described in this study encompassed an interactive design by encouraging self-monitoring and providing automated feedback, yet high levels of attrition and long-term disengagement levels were evident. Incorporation of more individualized personal support rather than automated feedback may have helped engagement levels, particularly at the 6-month juncture. The majority of participants who took part in an interview suggested that the addition of more personalized interaction, by a phone call or email, rather than an automated machine providing standard feedback would be more motivating and help preserve their interest to keep using the program. Similar studies evaluating the effect of Web-based programs on weight loss have reported higher effect sizes and usage when face-to-face contact is incorporated into the intervention [4]. This would be in-line with the complete Imperative Health service; however, it was important to assess the Web-based program on its own to understand its specific contribution.

Incorporation of Web-based programs into traditional care pathways for weight loss has generally taken the approach of comparing standard health care to standard health care plus a Web-based program over a defined period of time. An alternative model of care that may be worth further investigation is to use Web-based programs for initiation of weight loss and then add in further interventions rather than using Web-based programs alongside other interventions from the outset. Addition of more interpersonal interventions at the later stage would perhaps encourage sustained behavior change, prevent attrition from the 6-month time point onward, and support the weight loss maintenance stage. Such a model has particular relevance for health care systems. For example, waiting lists to be seen by dieticians in the UK National Health Service can be many weeks; referral to use a Web-based program during this time would be a useful way of initiating weight loss and may be particularly appealing for patients who do not feel comfortable attending weight loss groups.

Secondary Outcomes

In terms of cardiovascular risk, between-group analyses demonstrated that the intervention group significantly improved their BMI and waist circumference at 3 months and 6 months and their total cholesterol and triglycerides from baseline to 3 months in comparison with the control group; however, these significant between-group changes were not sustained at 12 months. The Web-based program did provide a self-monitoring tool for tracking blood pressure, lipids, and lipoprotein levels, but intervention group participants did not avail of this part of the program. When this was discussed at the interview sessions, the majority of participants stated that the main reason they did not access this part of the program was because they were not able to have these risk factors measured regularly. This is a general disadvantage of Web-based programs that do encourage the monitoring of other health risk factors, but do not provide the means to conduct the measurements at the participant's own convenience.

It was evident that the intervention group adopted healthier behaviors specifically in the short term. The significant increase in time spent exercising moderately and above (>3.5 METs) and the decrease in energy intake observed in the intervention group in comparison to the control group was likely to be attributable to the self-monitoring components of the Web-based program. Usages of these self-monitoring features were also correlated with short-term weight change (baseline to 3 and 6 months). Physical activity levels were not sustained in the longer term and the number of weekly physical activity uploads notably decreased between 6 and 12 months. These findings are consistent with those already reported in the literature [39,40]. A systematic review examining the effects of self-monitoring diet, physical activity, and weight on weight loss [39] found a consistent and positive significant association between the frequency of the self-monitoring behaviors and weight loss compared to less frequent self-monitoring. It was also reported in this review that a gradual decline over time in adherence to self-monitoring weight management behaviors is common [39].

Strengths and Limitations

The strengths of this study included its robust study design, the objectively measured primary outcome, and the mixed-method research approaches (qualitative and quantitative) used throughout the evaluation process. Furthermore, this study included a true control group. The majority of studies in this area tend to use a minimal support group to boost recruitment and decrease attrition; however, this may attenuate the relationship between groups. This study was conducted within a real-life setting and participants were not provided with strict instructions as to how often they should use the program; therefore, making the results more generalizable to overweight populations accessing these Web-based programs at home for their own self-care.

XSL•FO
RenderX

This study did have some limitations, for example, all participants had contact with the researcher during clinical assessments and knew this was a weight loss study, which in itself may have triggered a behavior change response and the "Hawthorne effect" [41] appears to be evident within the control group. The researcher, however, did not give any advice during the assessment period to either group.

Issues of attrition or loss to follow-up and nonusage attrition steadily increased over time, but this phenomenon is commonly reported in the literature in relation to weight loss management [35] and is not unique to Web-based programs. From a scientific perspective, attrition and nonusage attrition can impact on the likelihood of detecting a difference between groups when evaluating the treatments over longer periods of time; from a clinical perspective, it highlights the challenge of maintaining interest, motivation, and weight loss in the medium to long term. The increased attrition over the 12-month intervention period diminished the power of the study to detect a difference in change in weight loss and other endpoints between the intervention and control group in the longer term. However, the difference in weight change between the 2 groups at 12 months was very small and did not provide support for the Web-based program having a clinically significant advantage for long-term weight loss relative to a true control (usual care).

Owing to the routes of recruitment and the fact that people volunteered themselves for this study, the majority of the sample was from a higher social economic background. This could have potentially affected levels of engagement and attrition. This is not unique to this study, but does suggest the sample is not likely to be entirely representative of the general overweight and obese population. This will need to be borne in mind when considering the potential wider- or larger-scale impact of the Web-based behavior change intervention.

Conclusions

This study provides evidence that this Web-based program can be used to initiate clinically relevant weight loss of 5% or more and promote improvements in total cholesterol and triglyceride concentrations in the short term (3-6 months) in comparison with usual care. However, these changes were not sustained in the longer term (up to 12 months) and this appeared to correspond with a general decline in usage of the Web-based program over time. The fact that the study was powered on weight loss at the 3-month juncture and the high attrition rates at the 12-month time point in the intervention group, could have also prevented significant differences between the groups being identified, specifically at the later time point. Nevertheless, results of this study highlight a need to augment Web-based programs with further interventions after 6 months of usage, such as phone, email, or face-to-face support, to enhance engagement, prevent relapses, and encourage maintenance of weight loss in the longer term. The effectiveness and cost-effectiveness of such a model of weight management is worth further exploration

Acknowledgments

This study was supported by a PhD studentship provided by the Northern Ireland Department for Employment and Learning. Imperative Health (AXA PPP Healthcare Ltd) provided the Web-based program and monitoring devices without charge. The authors would like to acknowledge the Imperative Health team; specifically, John Crudgington and Chris Tomkins for their contribution, assistance, and support throughout this project. We also acknowledge Dr Caroline Mercer for undertaking the serum lipid and CRP analysis. We would like to thank all the participating volunteers for giving up their time and for their cooperation.

Authors' Contributions

MCM, JVW, and LJW were all involved in the conception of this study. All authors listed (SW, JVW, LJW, SJH, AM, ISY, CRC, KMA, and MCM) were involved in designing this research. SW with the assistance of AM conducted the research. SW with the guidance and assistance of CRC analyzed the data. SW wrote the manuscript and all authors were involved in reading, revising it critically, editing, and approving the final manuscript.

Conflicts of Interest

LW was involved in developing the nutrition and behavior change elements of the Web-based program. All other authors listed on the manuscript have no conflicts to declare.

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.2 [10].

[PDF File (Adobe PDF File), 154KB - jmir_v17i7e177_app1.pdf]

References

1. Ng M, Fleming T, Robinson M, Thomson B, Graetz N, Margono C, et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2014 Aug;384(9945):766-781. [doi: 10.1016/S0140-6736(14)60460-8] [Medline: 24880830]

XSL•FO
RenderX

2. 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 Jun;8(2):e10 [FREE Full text] [doi: 10.2196/jmir.8.2.e10] [Medline: 16867965]
3. Bennett GG, Glasgow RE. The delivery of public health interventions via the Internet: actualizing their potential. Annu Rev Public Health 2009 Apr;30:273-292. [doi: 10.1146/annurev.publhealth.031308.100235] [Medline: 19296777]
4. Kohl LF, Crutzen R, de Vries NK. Online prevention aimed at lifestyle behaviors: a systematic review of reviews. J Med Internet Res 2013 Jul;15(7):e146 [FREE Full text] [doi: 10.2196/jmir.2665] [Medline: 23859884]
5. 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]
6. Arem H, Irwin M. A review of web-based weight loss interventions in adults. Obes Rev 2011 May;12(5):e236-e243 [FREE Full text] [doi: 10.1111/j.1467-789X.2010.00787.x] [Medline: 20804523]
7. Manzoni GM, Pagnini F, Corti S, Molinari E, Castelnuovo G. Internet-based behavioral interventions for obesity: an updated systematic review. Clin Pract Epidemiol Ment Health 2011 Mar;7:19-28 [FREE Full text] [doi: 10.2174/1745017901107010019] [Medline: 21552423]
8. Kodama S, Saito K, Tanaka S, Horikawa C, Fujiwara K, Hirasawa R, et al. Effect of Web-based lifestyle modification on weight control: a meta-analysis. Int J Obes (Lond) 2012 May;36(5):675-685. [doi: 10.1038/ijo.2011.121] [Medline: 21694698]
9. Wieland LS, Falzon L, Sciamanna CN, Trudeau KJ, Brodney S, Schwartz JE, et al. Interactive computer-based interventions for weight loss or weight maintenance in overweight or obese people. Cochrane Database Syst Rev 2012 Aug;8:CD007675 [FREE Full text] [doi: 10.1002/14651858.CD007675.pub2] [Medline: 22895964]
10. Eysenbach G. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011 Dec;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
11. National Health Service. The General Practice Physical Activity Questionnaire: A Screening Tool to Assess Adult Physical Activity Levels, Within Primary Care. 2009 May. URL: http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/@ps/documents/digitalasset/dh_112134.pdf [accessed 2014-08-17] [WebCite Cache ID 6RtwwLqKt]
12. 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]
13. Tate DF, Jackvony EH, Wing RR. A randomized trial comparing human e-mail counseling, computer-automated tailored counseling, and no counseling in an Internet weight loss program. Arch Intern Med 2006 Aug;166(15):1620-1625. [doi: 10.1001/archinte.166.15.1620] [Medline: 16908795]
14. Polzien KM, Jakicic JM, Tate DF, Otto AD. The efficacy of a technology-based system in a short-term behavioral weight loss intervention. Obesity (Silver Spring) 2007 Apr;15(4):825-830. [doi: 10.1038/oby.2007.584] [Medline: 17426316]
15. Bennett GG, Herring SJ, Puleo E, Stein EK, Emmons KM, Gillman MW. Web-based weight loss in primary care: a randomized controlled trial. Obesity (Silver Spring) 2010 Feb;18(2):308-313 [FREE Full text] [doi: 10.1038/oby.2009.242] [Medline: 19696764]
16. Hurling R, Fairley BW, Dias MB. Internet-based exercise intervention systems: Are more interactive designs better? Psychology & Health 2006 Dec;21(6):757-772. [doi: 10.1080/14768320600603257]
17. Hurling R, Catt M, Boni MD, Fairley BW, Hurst T, Murray P, et al. Using internet and mobile phone technology to deliver an automated physical activity program: randomized controlled trial. J Med Internet Res 2007 Apr;9(2):e7 [FREE Full text] [doi: 10.2196/jmir.9.2.e7] [Medline: 17478409]
18. 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 Dec;10(4):e56 [FREE Full text] [doi: 10.2196/jmir.1108] [Medline: 19117828]
19. Imperative: Good Habits for Life. Improve your health for good URL: http://www.imperativehealth.com/ [accessed 2014-08-19] [WebCite Cache ID 6RwhzAalk]
20. Wing RR, Hamman RF, Bray GA, Delahanty L, Edelstein SL, Hill JO, Diabetes Prevention Program Research Group. Achieving weight and activity goals among diabetes prevention program lifestyle participants. Obes Res 2004 Sep;12(9):1426-1434 [FREE Full text] [doi: 10.1038/oby.2004.179] [Medline: 15483207]
21. Burke BS. The dietary history as a tool in research. J AM Diet Assoc 1947;23:1041-1046.
22. Block G. A review of validations of dietary assessment methods. Am J Epidemiol 1982 Apr;115(4):492-505. [Medline: 7041631]
23. Crawley H, Mills A, Patel S, Food Standards Agency. Food Portion Sizes, 3rd edition. London: The Stationary Office; 2002.
24. Besson H, Brage S, Jakes RW, Ekelund U, Wareham NJ. Estimating physical activity energy expenditure, sedentary time, and physical activity intensity by self-report in adults. Am J Clin Nutr 2010 Jan;91(1):106-114 [FREE Full text] [doi: 10.3945/ajcn.2009.28432] [Medline: 19889820]
25. Office for National Statistics. SOC2010 volume 3: the National Statistics Socio-economic classification (NS-SEC rebased on SOC2010) URL: http://www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/soc2010/

XSL•FO
RenderX

soc2010-volume-3-ns-sec--rebased-on-soc2010--user-manual/index.html [accessed 2015-03-20] [WebCite Cache ID 6XAHNPcHF]
26. Crabtree BF, Miller WL. Using codes and code manuals: a template organizing style of interpretation. In: Doing Qualitative Research. Thousand Oaks, CA: Sage Publications; 1999:163-178.
27. Vickers AJ, Altman DG. Statistics notes: analysing controlled trials with baseline and follow up measurements. BMJ 2001 Nov 10;323(7321):1123-1124 [FREE Full text] [Medline: 11701584]
28. Ware JH. Interpreting incomplete data in studies of diet and weight loss. N Engl J Med 2003 May 22;348(21):2136-2137. [doi: 10.1056/NEJMe030054] [Medline: 12761370]
29. World Health Organization. Obesity: Preventing and Managing the Global Epidemic. Geneva: World Health Organization; 2000.
30. 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]
31. McConnon A, Kirk SF, Cockroft JE, Harvey EL, Greenwood DC, Thomas JD, et al. The Internet for weight control in an obese sample: results of a randomised controlled trial. BMC Health Serv Res 2007 Dec;7:206 [FREE Full text] [doi: 10.1186/1472-6963-7-206] [Medline: 18093289]
32. Hunter CM, Peterson AL, Alvarez LM, Poston WC, Brundige AR, Haddock CK, Van Brunt David L, et al. Weight management using the internet a randomized controlled trial. Am J Prev Med 2008 Feb;34(2):119-126. [doi: 10.1016/j.amepre.2007.09.026] [Medline: 18201641]
33. Morgan PJ, Lubans DR, Collins CE, Warren JM, Callister R. The SHED-IT randomized controlled trial: evaluation of an Internet-based weight-loss program for men. Obesity (Silver Spring) 2009 Nov;17(11):2025-2032. [doi: 10.1038/oby.2009.85] [Medline: 19343018]
34. Morgan PJ, Lubans DR, Collins CE, Warren JM, Callister R. 12-month outcomes and process evaluation of the SHED-IT RCT: an internet-based weight loss program targeting men. Obesity (Silver Spring) 2011 Jan;19(1):142-151. [doi: 10.1038/oby.2010.119] [Medline: 20523304]
35. Moroshko I, Brennan L, O'Brien P. Predictors of dropout in weight loss interventions: a systematic review of the literature. Obes Rev 2011 Nov;12(11):912-934. [doi: 10.1111/j.1467-789X.2011.00915.x] [Medline: 21815990]
36. Carr LJ, Bartee RT, Dorozynski C, Broomfield JF, Smith ML, Smith DT. Internet-delivered behavior change program increases physical activity and improves cardiometabolic disease risk factors in sedentary adults: results of a randomized controlled trial. Prev Med 2008 May;46(5):431-438. [doi: 10.1016/j.ypmed.2007.12.005] [Medline: 18207228]
37. Cussler EC, Teixeira PJ, Going SB, Houtkooper LB, Metcalfe LL, Blew RM, et al. Maintenance of weight loss in overweight middle-aged women through the Internet. Obesity (Silver Spring) 2008 May;16(5):1052-1060. [doi: 10.1038/oby.2008.19] [Medline: 18309301]
38. Collins CE, Morgan PJ, Jones P, Fletcher K, Martin J, Aguiar EJ, et al. A 12-week commercial web-based weight-loss program for overweight and obese adults: randomized controlled trial comparing basic versus enhanced features. J Med Internet Res 2012 Apr;14(2):e57 [FREE Full text] [doi: 10.2196/jmir.1980] [Medline: 22555246]
39. Burke LE, Wang J, Sevick MA. Self-monitoring in weight loss: a systematic review of the literature. J Am Diet Assoc 2011 Jan;111(1):92-102 [FREE Full text] [doi: 10.1016/j.jada.2010.10.008] [Medline: 21185970]
40. Krukowski RA, Harvey-Berino J, Bursac Z, Ashikaga T, West DS. Patterns of success: online self-monitoring in a web-based behavioral weight control program. Health Psychol 2013 Feb;32(2):164-170. [doi: 10.1037/a0028135] [Medline: 22545978]
41. Parsons HM. What happened at Hawthorne?: New evidence suggests the Hawthorne effect resulted from operant reinforcement contingencies. Science 1974 Mar;183(4128):922-932. [doi: 10.1126/science.183.4128.922] [Medline: 17756742]

Abbreviations

BMI: body mass index
BOCF: baseline observation carried forward
CRP: C-reactive protein
CVD: cardiovascular disease
GPPAQ: General Practice Physical Activity Questionnaire
HDL: high-density lipoprotein
ITT: intention-to-treat
MET: Metabolic Equivalent Task
NE-SEC: National Statistics Socio-Economic Classification
QUB: Queen's University Belfast
RCT: randomized controlled trial
RPAQ: Recent Physical Activity Questionnaire

XSL•FO
RenderX

Edited by G Eysenbach; submitted 04.09.14; peer-reviewed by J Nunez-Cordoba, M Binks; comments to author 06.11.14; revised version received 14.12.14; accepted 27.04.15; published 16.07.15.

Please cite as:
Watson S, Woodside JV, Ware LJ, Hunter SJ, McGrath A, Cardwell CR, Appleton KM, Young IS, McKinley MC
Effect of a Web-Based Behavior Change Program on Weight Loss and Cardiovascular Risk Factors in Overweight and Obese Adults at High Risk of Developing Cardiovascular Disease: Randomized Controlled Trial
J Med Internet Res 2015;17(7):e177
URL: http://www.jmir.org/2015/7/e177/
doi:10.2196/jmir.3828
PMID:26183659

©Sinead Watson, Jayne V Woodside, Lisa J Ware, Steven J Hunter, Alanna McGrath, Christopher R Cardwell, Katherine M Appleton, Ian S Young, Michelle C McKinley. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 16.07.2015. 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 Mobile Insulin Titration Intervention (MITI) for Insulin Adjustment in an Urban, Low-Income Population: Randomized Controlled Trial

Natalie Levy[1,2], MD; Victoria Moynihan[1,2], BA; Annielyn Nilo[2], BSN, RN, CDE; Karyn Singer[3], MD, MPH; Lidia S Bernik[4], MHS, MBA; Mary-Ann Etiebet[2], MD, MBA; Yixin Fang[5], PhD; James Cho[1,2], MD; Sundar Natarajan[1,6,7], MD, MSc

[1]Division of General Internal Medicine and Clinical Innovation, Department of Medicine, New York University School of Medicine, New York, NY, United States

[2]Bellevue Hospital Center, New York, NY, United States

[3]Gouverneur Diagnostic and Treatment Center, New York, NY, United States

[4]Department of Healthcare Policy and Research, Weill Cornell Medical College, New York, NY, United States

[5]Division of Biostatistics, Department of Population Health, New York University School of Medicine, New York, NY, United States

[6]Department of Population Health, New York University School of Medicine, New York, NY, United States

[7]Department of Veterans Affairs, New York Harbor Healthcare System, New York, NY, United States

Corresponding Author:
Victoria Moynihan, BA
Division of General Internal Medicine and Clinical Innovation
Department of Medicine
New York University School of Medicine
Bellevue CD building, 630
462 First Ave
New York, NY, 10016
United States
Phone: 1 646 501 0691
Fax: 1 212 263 8788
Email: Victoria.Ramsay@nyumc.org

Abstract

Background: Diabetes patients are usually started on a low dose of insulin and their dose is adjusted or “titrated” according to their blood glucose levels. Insulin titration administered through face-to-face visits with a clinician can be time consuming and logistically burdensome for patients, especially those of low socioeconomic status (SES). Given the wide use of mobile phones among this population, there is the potential to use short message service (SMS) text messaging and phone calls to perform insulin titration remotely.

Objective: The goals of this pilot study were to (1) evaluate if our Mobile Insulin Titration Intervention (MITI) intervention using text messaging and phone calls was effective in helping patients reach their optimal insulin glargine dose within 12 weeks, (2) assess the feasibility of the intervention within our clinic setting and patient population, (3) collect data on the cost savings associated with the intervention, and (4) measure patient satisfaction with the intervention.

Methods: This was a pilot study evaluating an intervention for patients requiring insulin glargine titration in the outpatient medical clinic of Bellevue Hospital Center in New York City. Patients in the intervention arm received weekday SMS text messages from a health management platform requesting their fasting blood glucose values. The clinic’s diabetes nurse educator monitored the texted responses on the platform website each weekday for alarm values. Once a week, the nurse reviewed the glucose values, consulted the MITI titration algorithm, and called patients to adjust their insulin dose. Patients in the usual care arm continued to receive their standard clinic care for insulin titration. The primary outcome was whether a patient reached his/her optimal insulin glargine dose within 12 weeks.

Results: A total of 61 patients consented and were randomized into the study. A significantly greater proportion of patients in the intervention arm reached their optimal insulin glargine dose than patients in the usual care arm (88%, 29/33 vs 37%, 10/27; $P<.001$). Patients responded to 84.3% (420/498) of the SMS text messages requesting their blood glucose values. The nurse

XSL•FO
RenderX

reached patients within 2 attempts or by voicemail 91% of the time (90/99 assigned calls). When patients traveled to the clinic, they spent a median of 45 minutes (IQR 30-60) on travel and 39 minutes (IQR 30-64) waiting prior to appointments. A total of 61% (37/61) of patients had appointment copays. After participating in the study, patients in the intervention arm reported higher treatment satisfaction than those in the usual care arm.

Conclusions: MITI is an effective way to help low-SES patients reach their optimal insulin glargine dose using basic SMS text messaging and phone calls. The intervention was feasible and patients were highly satisfied with their treatment. The intervention was cost saving in terms of time for patients, who were able to have their insulin titrated without multiple clinic appointments. Similar interventions should be explored to improve care for low-SES patients managing chronic disease.

Trial Registration: Clinicaltrials.gov NCT01879579; https://clinicaltrials.gov/ct2/show/NCT01879579 (Archived by WebCite at http://www.webcitation.org/6YZik33L3).

(J Med Internet Res 2015;17(7):e180) doi:10.2196/jmir.4716

KEYWORDS

patient-centered care; health care disparities; telemedicine; remote consultation; cell phones; insulin, long-acting; text messaging

Introduction

Background

Many patients with diabetes mellitus in the United States are poorly controlled (glycated hemoglobin A_{1c} [HbA_{1c}] ≥9%). This includes 48.7% of the diabetics insured by Medicaid and 27.3% of diabetics insured by Medicare [1]. The consequences of uncontrolled diabetes are severe (eg, stroke, blindness, kidney disease, and amputation) and disproportionately affect patients of low socioeconomic status (SES) [2,3]. Insulin is commonly prescribed to treat uncontrolled diabetes [4]. Patients are started on a low dose of insulin and their dose is adjusted or "titrated" according to their blood glucose levels. Adjustments are made until the patient reaches a dose that best controls their glucose levels. Insulin titration traditionally occurs during a face-to-face encounter with a clinician [5-9]. Patients show the clinician their blood glucose levels from at-home testing and then the clinician recommends an appropriate insulin titration. One titration is often not enough to achieve glycemic control, so patients need to return to the clinic for multiple appointments.

Attending multiple appointments can be challenging for low-SES patients. They are faced with competing priorities that can make the many self-care tasks of diabetes management overwhelming [10-14]. Attending a clinic appointment can mean missing work hours with an inflexible job, lost wages, copays, and arranging for childcare and transportation to the clinic. Given these challenges, the process of insulin titration and achieving glycemic control may be prolonged.

Mobile phones are increasingly used to deliver health services [15]. Research shows 84% of the low-income US population owns a mobile phone, but only 47% of this population owns a phone with advanced features (ie, smartphone) [16]. Potential health interventions designed around basic features (eg, texting and voice) would not require a smartphone and, therefore, be the most accessible for low-income populations. In addition, these technologies still allow patients and clinicians to consult one another directly, allowing for personalized, nuanced care. A recent study of smartphone apps with insulin dose calculators showed that most have significant shortcomings. These apps may not take into account the patient's level of clinical knowledge, missing glucose readings, or concurrent oral antihyperglycemic medications, potentially introducing a safety risk [17].

Prior Research

Studies show that short message service (SMS) text messaging is an effective medium to assist with diabetes management in general and low-SES populations [6,18-25]. It can be used successfully to remind patients to check their blood glucose levels and to gather that data so that a clinician can review it for the next in-person clinic appointment [19].

Of the few studies in which clinicians titrated insulin remotely, the interventions typically required Internet access or website navigation. Patients sent their blood glucose values to their clinicians via the Internet. Clinicians responded to these data by sending their recommendations over the Internet or by SMS text message. These studies show that it is feasible to have patients send their blood glucose data and have clinicians relay insulin dose titration advice remotely [6,22,26]. However, with our intervention we aim to show that this exchange of data can be achieved using only basic text message and voice technology.

Current Intervention

The Mobile Insulin Titration Intervention (MITI) is a randomized controlled trial for patients who require insulin glargine titration. We chose to focus on glargine because it is the type of insulin used in our hospital formulary. Our intervention uses features available on basic mobile phones: SMS text messaging and voice calls. This technology is easy to use, low-cost, and widely available to our patient population. Through a text message, we can remind patients to check their glucose at any time and place that they have phone service. Patients can respond via text quickly and simply. Using weekly phone calls, patients and clinicians can still discuss their insulin treatment in a personal manner without the burden of an in-person appointment.

Through the MITI study, we aimed to (1) determine if MITI is effective in helping patients reach their optimal dose of insulin glargine ("optimal dose" is defined in the Intervention section), (2) evaluate the feasibility of the intervention, (3) measure the cost savings associated with the intervention, and (4) measure patient satisfaction.

XSL•FO
RenderX

Methods

Study Design

Overview

The MITI study is a randomized controlled trial that uses a parallel study design. Patients randomized to the MITI arm were allocated to receive the intervention. Patients randomized to the usual care arm acted as the control group and continued to receive standard care in the clinic. Patient outcomes were tracked for 12 weeks after enrollment in the study (see Outcome Measures section). A data and safety monitoring board reviewed any potential safety concerns for the duration of the study. This paper summarizes the methods for this trial. Further details are available in the published protocol [27].

Setting

This study occurred at Bellevue Hospital Center in New York City. Bellevue is part of the Health and Hospitals Corporation (HHC), the largest public hospital system in the United States. We recruited patients for this study from Bellevue's Adult Primary Care Center (APCC). Most clinic visits are for patients who are uninsured (31%) or have Medicaid (45%) [28]. The majority of patients are nonwhite: Hispanic (41%), Black (24%), and Asian (6%) (Bellevue Hospital Center, unpublished data, 2014). The prevalence rate of diabetes in the APCC is 15% (HHC Patient Registry for Proactive Care, unpublished data, 2014), higher than the national rate of 9.3% [2].

Inclusion and Exclusion Criteria

The inclusion criteria for patients were initiating insulin glargine or requiring the titration of an existing insulin glargine dose, English or Spanish speaking, the most recent HbA_{1c} value at or above 8%, able and willing to inject insulin, and able and willing to provide informed consent. The exclusion criteria were patients on short-acting insulin, on systemic glucocorticoids, with sustained serum creatinine at or above 1.5 mg/dL for men and 1.4 mg/dL for women, with documented hypoglycemia unawareness, and with type 1 diabetes.

Recruitment and Enrollment

Patients were recruited for the study from the APCC. Clinicians referred patients to the research assistant (RA). The RA screened patients for eligibility and enrolled them in the study. The enrollment process occurred in-person in the clinic and all patients provided informed consent before participating in the study. Patients were given a US $10 Metrocard for transportation to/from the clinic and an additional US $10 Metrocard at follow-up. Patients were also given a blood glucose meter and test strips to allow them to check their blood glucose levels during the study.

Randomization

Patients were randomized on the day of enrollment after the informed consent process. The random allocation sequence was computer-generated by a coinvestigator and concealed in sequentially numbered envelopes. Patients were stratified by whether they were initiating insulin treatment or having their existing insulin dose adjusted. Within each stratification, the allocation sequence used blocks of 4 to help keep the number of patients balanced in each arm. Patients, clinicians, and researchers in this trial were not blinded to arm assignments.

Intervention

After consent and randomization, patients in the MITI arm enrolled in a Web-based health management platform during the enrollment process at the clinic. The platform automatically sent patients a text message each weekday morning asking them for their fasting blood glucose value. During enrollment, the patient was able to choose English or Spanish messages and the specific time of day when the messages would be sent. When patients received the text message on their phone, they responded with their blood glucose value. The diabetes nurse educator logged onto her secure account on the platform each weekday afternoon to view the patients' text message responses. She would call any patient that had texted an alarm value (blood glucose <80 or >400 mg/dL). Patients were instructed to call the diabetes nurse educator (which is the standard practice with patients in the clinic) in addition to sending the text if they had an alarm value.

Each Thursday afternoon, the nurse reviewed the texted values, consulted the titration algorithm (which was developed by physicians and nurses on the study team), and called the patient to adjust his/her insulin dose. The nurse could call the patient's emergency contact on her discretion. Beginning in May 2014, we revised our study protocol and outlined voicemail as an option for the nurse to give titration instructions to patients. When the nurse was not available to check the text responses for alarm values or to make titration calls, a physician on the study team performed this task.

Patients continued with the weekday SMS text messages and weekly phone calls until they reached their optimal insulin glargine dose or for a maximum of 12 weeks. We defined optimal insulin dose as the dose at which a patient achieved at least 1 fasting blood glucose value between 80 and 130 mg/dL (inclusive) or the maximum dose that could be safely administered to the patient. During the intervention, patients continued to attend appointments with their primary care provider, but did not need to attend appointments specifically for diabetes management (eg, high HbA_{1c} clinic or diabetes nurse educator appointments). After completing the intervention, patients resumed usual clinic care. The study team arranged any follow-up appointments needed to allow the patient to resume their standard diabetes care (eg, primary care provider, diabetes nurse educator, and high HbA_{1c} clinic appointments).

Usual Care

After consent and randomization, patients were instructed to continue with their existing treatment plan and appointments for diabetes care. After the patient had a clinic appointment for insulin titration, the RA collected data (in-person or by phone) about the appointment. These data included the patient's insulin dose, blood glucose values, and data for cost savings outcomes.

Follow-Up

At approximately 12 weeks after study enrollment, patients in both arms were contacted by the RA to remind them of their

XSL•FO
RenderX

routine HbA_{1c} test and to ask them to fill out the Diabetes Treatment Satisfaction Questionnaire (either over the phone or when the patient was in the clinic).

Implementation Challenges

Our initial health management platform was not able to send SMS text messages to patients with prepaid phones; thus, these patients were not able to sign up to participate in the intervention (see Participant Characteristics). These patients continued to attend in-person appointments for insulin titration and were included in the intention-to-treat analysis. Subsequent patients with incompatible phones were provided a mobile phone to use during the study. Beginning in May 2014, patients were enrolled using a different health management platform that accommodated all types of mobile phones.

We initially stratified participants by insulin treatment status (initiating insulin or needing their existing dose titrated) and by HbA_{1c} level (8%-11% or >11%). In May 2014, we decided to stratify only by insulin treatment status after finding that not all participants had an HbA_{1c} value in their medical record on the day of study enrollment.

Outcome Measures

Our primary outcome was whether a patient reached his/her optimal insulin glargine dose within 12 weeks of enrolling in the study. We hypothesized that a greater proportion of patients in the MITI arm would reach their optimal insulin dose as compared to the usual care arm. The research staff recorded whether a patient reached his/her optimal insulin dose after each titration phone call (for the MITI arm) or after each clinic appointment (for the usual care arm). We also examined the time it took to reach optimal dose, patient self-reported hypoglycemia, and change in HbA_{1c} levels between baseline and 12 weeks.

We measured the feasibility of the intervention, including patient text response rate, the ability of the nurse to reach patients for titration phone calls, and the time the nurse spent on the intervention.

We collected data on the cost savings associated with the intervention. These data included the number and duration of insulin titration interactions (appointments during which insulin was titrated), the time patients spent traveling to the clinic and waiting prior to appointments, copays for clinic appointments, and patient health care utilization (the number of walk-in clinic, medication refill, and emergency room visits at Bellevue Hospital Center). Copays refer to the amount that the patient pays the clinic on attending an appointment with a health care provider. For patients with insurance plans, the amount is typically set by the insurance company. For uninsured patients at our hospital, the amount is based on income. For patients in our study, the most common copay was US $15.

To assess patient satisfaction with the intervention, we used the Diabetes Treatment Satisfaction Questionnaire (status version) [29]. This was administered at study enrollment and approximately 12 weeks later. We also administered the Diabetes Treatment Satisfaction Questionnaire (change version) [29] to measure the change in satisfaction after study participation. Patients in the MITI arm participated in a semistructured interview to give qualitative feedback on the intervention. This occurred when the patient reached his/her optimal dose or when the 12 weeks had elapsed.

Statistical Analysis

Baseline characteristics were summarized using descriptive statistics and compared to determine if the arms were balanced. Chi-square tests were used for categorical outcomes and Wilcoxon rank sum tests for continuous outcomes. Interval-censoring survival analysis was used to analyze the time to reach optimal insulin dose. The generalized estimating equation (GEE) modeling was used for repeatedly measured text message responses and the duration of titration interactions. Multiple imputation was used to deal with missing data in HbA_{1c} measures. Intention-to-treat analysis was used.

Results

Participant Characteristics

Patients were recruited from June 2013 to December 2013 and May 2014 to December 2014. Follow-up data were collected until March 2015. We screened 132 patients for eligibility; 54 were ineligible and 17 declined to participate (Figure 1). A total of 61 patients consented and were randomized into the study; 33 in the MITI arm and 28 in the usual care arm. There were 36 patients who were stratified as new to insulin treatment and 25 that were having their existing insulin dose adjusted. Of these 61 patients, there were 6 patients (5 in MITI and 1 in usual care) who met inclusion criteria when screened at the time of enrollment, but were discovered to be ineligible to participate soon after they consented and were randomized. Of the 5 ineligible patients randomized to the MITI arm, 3 had prepaid mobile phones that were not able to sign up for our SMS text messaging platform, 1 was not starting insulin glargine, and 1 did not return to the clinic to complete enrollment. The ineligible patient randomized to the usual care arm phenotypically fit a type 1 diabetes diagnosis. These 6 patients did not receive the allocated intervention, but were included in the intention-to-treat analysis. No significant differences in baseline characteristics/demographics were found between the 2 study arms. Demographics of participants are shown in Table 1.

XSL•FO
RenderX

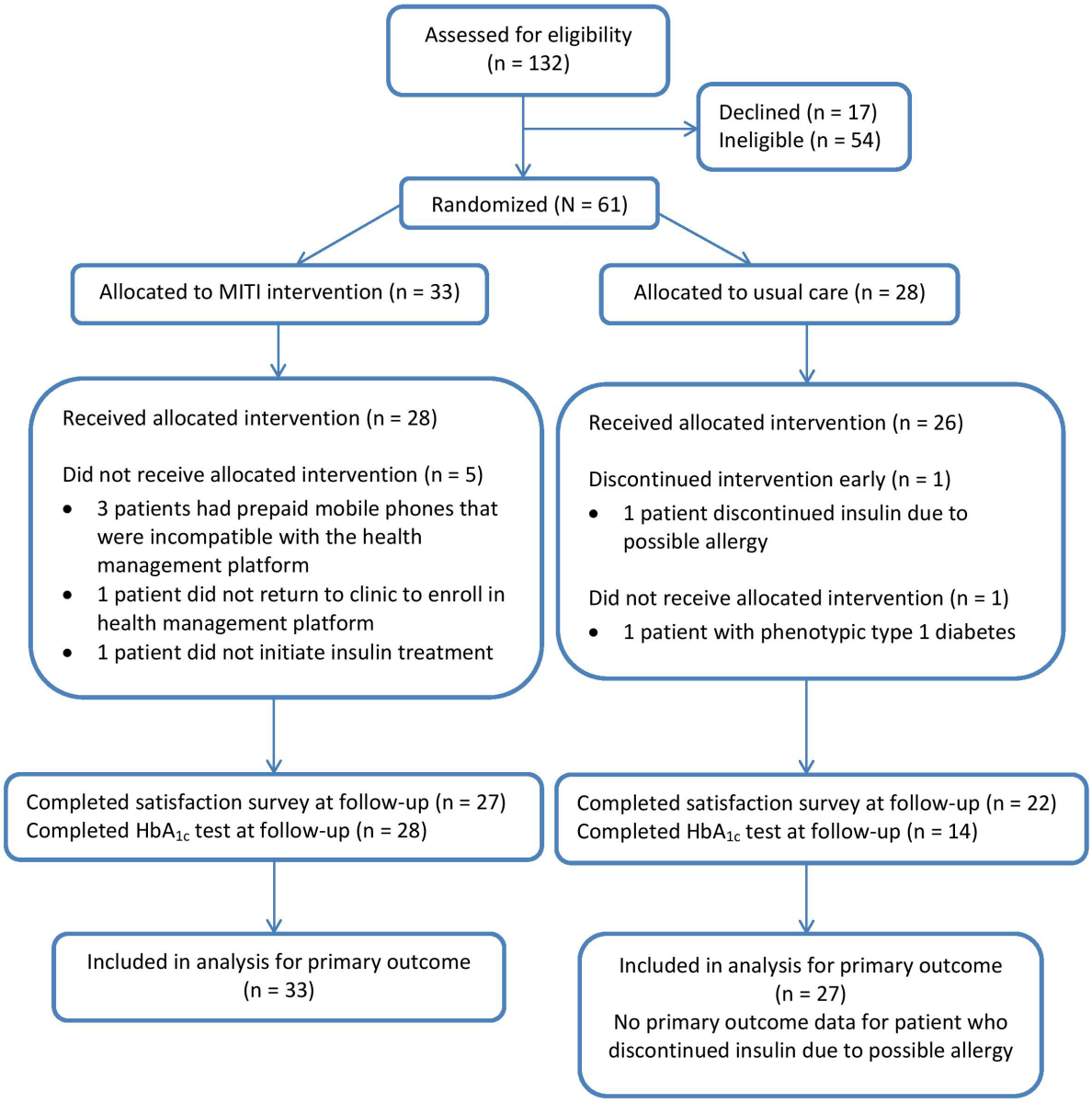
Table 1. Demographics of participants.

Demographics	MITI (n=33)	Usual care (n=28)	Total (N=61)	*P*
Age, mean (SD)	48.48 (11.22)	44.61 (9.97)	46.70 (10.75)	.14
Gender (female), n (%)	15 (45)	16 (57)	31 (51)	.36
Self-identified race/ethnicity, n (%)				.85
Hispanic	18 (55)	17 (61)	35 (57)	
Black or African American	8 (24)	7 (25)	15 (25)	
White	4 (12)	2 (7)	6 (10)	
Asian	2 (6)	2 (7)	4 (7)	
Caribbean	1 (3)	0 (0)	1 (2)	
Highest education level, [a] n (%)				.65
≤Grade 8	5 (16)	8 (29)	13 (22)	
Some high school	4 (13)	4 (14)	8 (13)	
High school graduate/GED	13 (41)	9 (32)	22 (37)	
Some college	4 (13)	3 (11)	7 (12)	
College degree	2 (6)	3 (11)	5 (8)	
Graduate degree	4 (13)	1 (4)	5 (8)	
Annual household income (US $), n (%)				.87
No response	18 (55)	9 (32)	27 (44)	
<10,000	3 (9)	5 (18)	8 (13)	
10,000-19,999	5 (15)	6 (21)	11 (18)	
20,000-29,999	2 (6)	4 (14)	6 (10)	
30,000-39,999	3 (9)	3 (11)	6 (10)	
>40,000	2 (6)	1 (4)	3 (5)	
Currently without health insurance, n (%)	23 (70)	15 (54)	38 (62)	.20
Baseline HbA_{1c}, mean (SD)	11.43 (1.75)	12.05 (1.91)	11.72 (1.83)	.14

[a] For MITI arm, n=32.

XSL•FO
RenderX

Figure 1. CONSORT diagram.



Clinical Outcomes

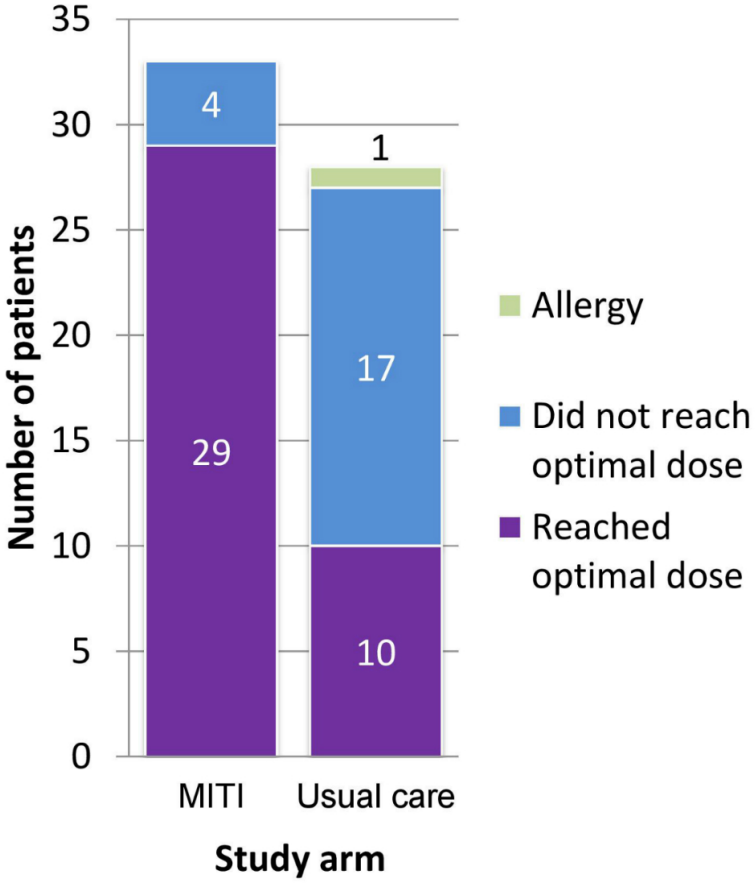
Primary Outcome: Reaching Optimal Insulin Dose

The primary outcome was the number of patients that reached their optimal insulin glargine dose within 12 weeks. In the MITI arm, 29 of 33 patients (88%, 95% CI 72%-97%) reached their optimal dose. Of the 29 patients who met optimal dose, 27 did so by achieving a fasting blood glucose value between 80 and 130 mg/dL (inclusive). Two patients reached the maximum dose that could be safely administered. In the usual care arm, 10 of 27 patients (37%, 95% CI 19%-58%) reached their optimal dose (Figure 2). Of the 10 patients in the usual care arm that reached their optimal dose, 9 did so by achieving a fasting blood glucose between 80 and 130 mg/dL. One patient met this goal by reaching the maximum dose that could safely be administered. The primary outcome could not be measured for one usual care patient who discontinued insulin glargine early due to a possible allergic reaction. The MITI arm had a significantly greater proportion of patients reach their optimal insulin glargine dose (OR 12.3, 95% CI 3.3-45.4, $P<.001$).

For the 29 patients in the MITI arm that reached their optimal insulin glargine dose, the median time to optimal dose was 3.00 weeks (IQR 1.29-4.86). For the 10 patients in the usual care arm that reached optimal dose, the median time was 7.07 weeks (IQR 2.96-9.61, $P=.007$).

XSL•FO
RenderX

Figure 2. Number of patients who reached optimal insulin dose within 12 weeks. Note: for patients that did not reach optimal dose, 3 of 4 patients in the MITI arm and 1 of 17 in the usual care arm did not receive the allocated intervention.



Glycated Hemoglobin

We measured HbA_{1c} change from baseline to 12 weeks. We included HbA_{1c} values from routine blood tests drawn within 4 weeks of baseline (study enrollment date) and 12 weeks after the baseline HbA_{1c} test (± 4 weeks). Looking at the nonimputed dataset, the mean HbA_{1c} for the MITI arm was 11.30% (SD 1.79, n=30) at baseline and 9.34% (SD 1.45, n=28) at 12 weeks. For the usual care arm, the mean was 12.20% (SD 1.90, n=25) at baseline and 9.99% (SD 1.33, n=14) at 12 weeks. The mean change for patients with an HbA_{1c} value at both baseline and 12 weeks was calculated. There were 28 patients in the MITI arm and 14 in the usual care arm. The mean change in HbA_{1c} between baseline and 12 weeks for the MITI arm was –1.90 (SD 2.64, n=28) and –1.81 (SD 2.63, n=14) for the usual care arm (*P*=.99). Combining the results from 10 multiple imputations (monotone method used), HbA_{1c} values in the MITI arm were 0.85 points lower than the usual care arm at 12 weeks (95% CI –1.83 to 0.13, *P*=.09). The large difference between the raw result and the multiple imputation result indicates that the missing mechanism was missing-not-at-random and the missing data problem was a limitation of this study for examining HbA_{1c} change.

Adverse Outcomes

There were 5 cases of hypoglycemia; 3 patients in the MITI arm and 2 in the usual care arm. All cases were mild with blood glucose values ranging between 69 and 79 mg/dL and none of the patients required assistance. One patient had a potential mild allergic reaction to insulin glargine.

Feasibility Outcomes

Text Message Responses

Of the 498 SMS text messages that were successfully sent to patients asking for their fasting blood glucose level, 420 (84.3%) received a reply. The GEE-adjusted response rate was 91.6% after adjusting for the difference in the number of text messages sent to/from patients and the correlation between responses. The mean number of SMS text messages successfully sent per patient was 18 texts (range 2-60). One patient did not receive the text prompts but was able to send back glucose values via text. The mean number of patient replies received was 16 texts (range 0-57). One patient who did receive the texts attempted to reply, but was not able to send the texts.

Titration Phone Calls

We reviewed each week that the nurse was assigned to call patients for their titration phone call (99 assigned calls total). The nurse was able to reach patients on the first or second attempt or by voicemail 91% of the time (90/99 assigned calls).

Time Spent by the Nurse on the Intervention

We reviewed the amount of time the diabetes nurse educator spent on the intervention. The vast majority of days (unless there was a technical issue that delayed Internet access), it took

XSL•FO
RenderX

the nurse 1 minute or less to review texted blood glucose values on the Web portal. On titration Thursdays, the nurse spent approximately 7-11 minutes per patient. When assigned to call one patient, she spent a mean total of 11.2 minutes (SD 7.4) making phone calls. She spent a mean total of 17.6 (SD 10.9) minutes with 2 patients, 26.5 minutes (SD 14.5) with 3 patients, 36.2 minutes (SD 7.5) with 4 patients, and 36.0 minutes (SD 9.0) with 5 patients.

Cost Savings Outcomes

We recorded the number of titration interactions (any interaction with a clinician to address insulin dosage either by phone/voicemail or in-person). The MITI arm had 131 interactions: 75.6% (99/131) by phone/voicemail and 24.4% (32/131) in-person. The usual care arm had 49 interactions: 98.0% (48/49) were in-person and 2.0% (1/49) by phone. Looking only at those patients who received the allocated intervention, the MITI arm had a median of 3.5 (IQR 2.0-5.0) titration interactions and usual care had 2.0 (IQR 1.0-3.0, *P*=.003). As expected, MITI patients had more titration interactions by phone than in-person. The median number of phone interactions was 3.0 (IQR 1.0-5.0) for MITI. Only one patient in the usual care arm had a phone interaction occur during the study. The median number of in-person titration interactions for MITI was 1.0 (IQR 1.0-1.0) and 2.0 (IQR 1.0-3.0) for usual care (*P*=.009).

We also recorded the duration of titration interactions and compared the duration by type of interaction (not by study arm). The median duration of titration interactions in the clinic was 30.0 minutes (IQR 20.0-45.0). The median for phone/voicemail interactions was 6.0 minutes (IQR 3.0-10.0). The difference in duration between clinic and phone/voicemail interactions was statistically significant (*P*=.008).

We looked at patient utilization of the Bellevue health care system for appointments other than insulin titration (walk-in, emergency department, and medication refill visits). The MITI arm had no increased utilization of the health care system over the 12 weeks.

We asked our study participants in both arms if they had copays for appointments at our clinic. Of 61 participants, 37 (61%) reported that they had copays. Of those participants, 32 (86%) reported a copay of US $15.

We asked study participants in both arms about their travel time and wait time when they visited the clinic. Participants reported a median travel time to the clinic (1-way) of 45 minutes (IQR 30-60). The median wait time prior to appointments was 39 minutes (IQR 30-64).

Satisfaction Outcomes

The Diabetes Treatment Satisfaction Questionnaire status version has 6 questions that assess patient satisfaction with diabetes treatment using a 0-6 scale (0 being very dissatisfied and 6 being very satisfied). At the time of enrollment, the mean score for the MITI arm was 4.99 (SD 1.14, n=32) and 5.20 (SD 0.61, n=28) for the usual care arm (*P*=.78). At approximately 12 weeks after enrollment, the mean for the MITI arm was 5.74 (SD 0.54, n=27) and 5.53 (SD 0.52, n=22) for the usual care arm (*P*=.04). The mean difference between the baseline and 12-week scores for the MITI arm (n=27) was 0.80 and 0.34 for the usual care arm (n=22, *P*=.16).

The Diabetes Treatment Satisfaction Questionnaire Change version has 6 questions that assess the change in patient satisfaction over the course of the study using a –3 to +3 scale (–3 being much less satisfied now and +3 being much more satisfied now). The mean change score for MITI was 2.71 (SD 0.71) and 2.42 (SD 0.95) for the usual care arm (*P*=.13).

Patients responded to the 12-week follow-up questionnaires within a mean of 1.4 weeks (SD 1.5) and median of 0.8 (IQR 0.0-2.0) weeks before or after the 12-week date. The follow-up time for the satisfaction questionnaires ranged from 3.1 weeks before to 6.1 weeks after the 12-week date.

A total of 27 patients in the MITI arm offered qualitative feedback on the intervention. Figure 3 shows selected quotes (from 11 patients) transcribed from the interview. The 14 quotes included in the figure are among the most specific comments made by patients about their experiences in the intervention.

Figure 3. Patient feedback on the intervention.



XSL•FO
RenderX

Discussion

Principal Findings

Clinical Outcomes

Our study showed that with simple SMS text messaging (requiring only basic mobile phone technology) and weekly titration phone calls, 88% of diabetic patients reach their optimal insulin glargine dose within 12 weeks (vs 37% in usual care, $P<.001$). This outcome was achieved without an increase in hypoglycemia.

Feasibility Outcomes

Our study sent a daily SMS text message asking for a reply text with the morning's fasting blood glucose and 84.3% of our texts received a reply. Our diabetes study nurse was able to relay titration instructions via phone calls for 91% of the assigned titration calls (within 2 tries or with a voicemail). Taken together, the response rates to our SMS text messages and the ease of reaching patients by phone support the feasibility of this intervention.

Cost Savings Outcomes

Patients enrolled in the MITI (texting) arm of our study saved travel time to and from the clinic, time spent in the waiting room, and the cost of copays, which 61% are required to pay at clinic appointments.

Satisfaction Outcomes

Through the Diabetes Treatment Satisfaction Questionnaire, we learned that MITI patients at 12 weeks were more satisfied with their treatment than usual care patients. The 14 quotes included in Figure 3 illustrate what some of our patients liked about their treatment during the intervention. These common themes include:

1. Text messages served as helpful reminders to check their glucose each morning.
2. Patients found the MITI intervention convenient.
3. Text messages made patients feel supported and cared for.
4. Patients liked knowing in a timely manner whether their insulin dose was adequate.
5. The need to check their blood glucose and report the values made patients feel more accountable for what they were eating and for taking their medications.

Limitations

The generalizability of this study is limited for several reasons. With limited manpower for patient recruitment and enrollment, we were not able to meet our target sample size of 49 patients per arm. Voluntary participants may not be representative of the clinic population as a whole. We do not know if the gains of the intervention (the motivation to be more compliant with diet, exercise, daily insulin use, home glucose monitoring, etc) lasted beyond the 12 weeks of the study. Missing data (50% of usual care patients did not have a 12-week HbA_{1c} test) was a limitation in examining change in HbA_{1c}. Patients must travel to the clinic to receive an HbA_{1c} blood test, which is a potential reason for the lack of HbA_{1c} data.

Conclusions

MITI was shown to be a highly effective way to titrate insulin glargine. We used daily automated SMS text messages (simple mobile phone technology, no app required) to gather fasting blood glucose values. We used weekly nurse phone calls (based on a physician-approved algorithm) to deliver titration advice. This intervention was well-suited to our urban clinic population. MITI was feasible, time saving, and satisfactory. MITI should be implemented with a larger sample of patients to further test its effectiveness. Similar intervention models should be explored for other diabetic medication titrations as well as other challenging aspects of chronic disease care.

Acknowledgments

We would like to thank our funders, the New York University-Health and Hospitals Corporation Clinical and Translational Science Institute (NYU-HHC CTSI) for the 2013 NYU CTSI Pilot Grant and the 2014 HHC H-3 Research Grant award # UL1 TR000038 from the National Center for the Advancement of Translational Science (NCATS), National Institutes of Health. We would also like to thank the HHC Office of Healthcare Improvement for their donation of diabetic testing supplies. We also acknowledge and thank Drs Patrick Cocks, Ming-Chin Yeh, Judith Wylie-Rosett, Ellie Grossman, Andrew Wallach, and Sondra Zabar for their guidance and support.

Authors' Contributions

NL, AN, VM, KS, LSB, M-AE, YF, JC, and SN made significant contributions to the planning and design of the study. NL, AN, VM, and JC implemented the study in the Adult Primary Care Center. YF performed the data analysis.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.2 [30].

[PDF File (Adobe PDF File), 147KB - jmir_v17i7e180_app1.pdf]

XSL•FO
RenderX

References

1. Health Resources and Services Administration. Diabetes HbA1c {Poor Control} URL: http://www.hrsa.gov/quality/toolbox/measures/diabetes/ [accessed 2015-05-16] [WebCite Cache ID 6YZnCxmxx]
2. American Diabetes Association. 2014 Jun 10. Statistics about diabetes: Data from the National Diabetes Statistics Report, 2014 (released June 10, 2014) URL: http://www.diabetes.org/diabetes-basics/statistics/ [accessed 2015-05-16] [WebCite Cache ID 6YZnJPa46]
3. World Health Organization. In: Blas E, Kurup AS, editors. Equity, Social Determinants and Public Health Programmes (Nonserial Publications). Geneva: World Health Organization; 2010.
4. Centers for Disease Control and Prevention. National Diabetes Statistics Report: Estimates of Diabetes and Its Burden in the United States, 2014. Atlanta, GA: US Department of Health and Human Services; 2014. URL: http://www.cdc.gov/diabetes/pubs/statsreport14/national-diabetes-report-web.pdf [accessed 2015-05-16] [WebCite Cache ID 6YZoGZMVV]
5. McCulloch DK. UpToDate. 2015. Patient information: diabetes mellitus type 2: insulin treatment (beyond the basics) URL: http://www.uptodate.com/contents/diabetes-mellitus-type-2-insulin-treatment-beyond-the-basics [accessed 2015-05-16] [WebCite Cache ID 6YZoUKqOP]
6. Vähätalo M, Virtamo H, Viikari J, Rönnemaa T. Cellular phone transferred self blood glucose monitoring: prerequisites for positive outcome. Pract Diab Int 2004 Jul 26;21(5):192-194. [doi: 10.1002/pdi.642]
7. Weidman-Evans E, Evans J, Eastwood R, Fort A. Implementation of a pharmacist-run telephonic insulin titration service. J Am Pharm Assoc (2003) 2012 Nov;52(6):e266-e272. [doi: 10.1331/JAPhA.2012.11225] [Medline: 23229989]
8. Bergenstal RM, Bashan E, McShane M, Johnson M, Hodish I. Can a tool that automates insulin titration be a key to diabetes management? Diabetes Technol Ther 2012 Aug;14(8):675-682 [FREE Full text] [doi: 10.1089/dia.2011.0303] [Medline: 22568777]
9. Khunti K, Davies MJ, Kalra S. Self-titration of insulin in the management of people with type 2 diabetes: a practical solution to improve management in primary care. Diabetes Obes Metab 2013 Aug;15(8):690-700. [doi: 10.1111/dom.12053] [Medline: 23253563]
10. Robert Wood Johnson Foundation. 2013 Aug 15. Low-income patients face added challenges in managing diabetes URL: http://www.rwjf.org/en/library/articles-and-news/2013/08/low-income-patients-face-added-challenges-in-managing-diabetes.html [accessed 2015-05-16] [WebCite Cache ID 6YZpDJJU9]
11. Gucciardi E, Demelo M, Offenheim A, Stewart DE. Factors contributing to attrition behavior in diabetes self-management programs: a mixed method approach. BMC Health Serv Res 2008 Feb;8:33 [FREE Full text] [doi: 10.1186/1472-6963-8-33] [Medline: 18248673]
12. Rosal MC, Heyden R, Mejilla R, Capelson R, Chalmers KA, Rizzo DM, et al. A virtual world versus face-to-face intervention format to promote diabetes self-management among African American women: a pilot randomized clinical trial. JMIR Res Protoc 2014 Oct;3(4):e54 [FREE Full text] [doi: 10.2196/resprot.3412] [Medline: 25344620]
13. Devoe JE, Baez A, Angier H, Krois L, Edlund C, Carney PA. Insurance + access not equal to health care: typology of barriers to health care access for low-income families. Ann Fam Med 2007 Dec;5(6):511-518 [FREE Full text] [doi: 10.1370/afm.748] [Medline: 18025488]
14. US Department of Health and Human Services, Health Resources and Services Administration. Women's Health USA 2011. Rockville, MD: US Department of Health and Human Services; 2011 Oct. URL: http://www.mchb.hrsa.gov/whusa11/more/downloads/pdf/w11.pdf [accessed 2015-05-16] [WebCite Cache ID 6YZpTe4OP]
15. Free C, Phillips G, Watson L, Galli L, Felix L, Edwards P, et al. The effectiveness of mobile-health technologies to improve health care service delivery processes: a systematic review and meta-analysis. PLoS Med 2013 Jan;10(1):e1001363 [FREE Full text] [doi: 10.1371/journal.pmed.1001363] [Medline: 23458994]
16. Mobile technology fact sheet. Washington, DC: Pew Internet & American Life Project; 2014 Oct. URL: http://www.pewinternet.org/fact-sheets/mobile-technology-fact-sheet/ [accessed 2014-12-31] [WebCite Cache ID 6VEeGRIkq]
17. Huckvale K, Adomaviciute S, Prieto JT, Leow MK, Car J. Smartphone apps for calculating insulin dose: a systematic assessment. BMC Med 2015;13:106 [FREE Full text] [doi: 10.1186/s12916-015-0314-7] [Medline: 25943590]
18. Arora S, Peters AL, Agy C, Menchine M. A mobile health intervention for inner city patients with poorly controlled diabetes: proof-of-concept of the TExT-MED program. Diabetes Technol Ther 2012 Jun;14(6):492-496. [doi: 10.1089/dia.2011.0252] [Medline: 22524591]
19. Fischer HH, Moore SL, Ginosar D, Davidson AJ, Rice-Peterson CM, Durfee MJ, et al. Care by cell phone: text messaging for chronic disease management. Am J Manag Care 2012 Feb;18(2):e42-e47 [FREE Full text] [Medline: 22435883]
20. Hussein WI, Hasan K, Jaradat AA. Effectiveness of mobile phone short message service on diabetes mellitus management; the SMS-DM study. Diabetes Res Clin Pract 2011 Oct;94(1):e24-e26. [doi: 10.1016/j.diabres.2011.07.025] [Medline: 21840079]
21. Dick JJ, Nundy S, Solomon MC, Bishop KN, Chin MH, Peek ME. Feasibility and usability of a text message-based program for diabetes self-management in an urban African-American population. J Diabetes Sci Technol 2011 Sep;5(5):1246-1254 [FREE Full text] [Medline: 22027326]
22. Kim HS, Jeong HS. A nurse short message service by cellular phone in type-2 diabetic patients for six months. J Clin Nurs 2007 Jun;16(6):1082-1087. [doi: 10.1111/j.1365-2702.2007.01698.x] [Medline: 17518883]

XSL•FO
RenderX

23. Osborn CY, Mulvaney SA. Development and feasibility of a text messaging and interactive voice response intervention for low-income, diverse adults with type 2 diabetes mellitus. J Diabetes Sci Technol 2013 May;7(3):612-622 [FREE Full text] [Medline: 23759393]
24. Nundy S, Dick JJ, Chou C, Nocon RS, Chin MH, Peek ME. Mobile phone diabetes project led to improved glycemic control and net savings for Chicago plan participants. Health Aff (Millwood) 2014 Feb;33(2):265-272 [FREE Full text] [doi: 10.1377/hlthaff.2013.0589] [Medline: 24493770]
25. Dobson R, Carter K, Cutfield R, Hulme A, Hulme R, McNamara C, et al. Diabetes text-message self-management support program (SMS4BG): a pilot study. JMIR Mhealth Uhealth 2015;3(1):e32 [FREE Full text] [doi: 10.2196/mhealth.3988] [Medline: 25830952]
26. Tildesley HD, Mazanderani AB, Ross SA. Effect of Internet therapeutic intervention on A1C levels in patients with type 2 diabetes treated with insulin. Diabetes Care 2010 Aug;33(8):1738-1740 [FREE Full text] [doi: 10.2337/dc09-2256] [Medline: 20668152]
27. Levy N, Moynihan V, Nilo A, Singer K, Bernik LS, Etiebet MA, et al. The Mobile Insulin Titration Intervention (MITI) for insulin glargine titration in an urban, low-income population: randomized controlled trial protocol. JMIR Res Protoc 2015;4(1):e31 [FREE Full text] [doi: 10.2196/resprot.4206] [Medline: 25794243]
28. Pressman M, Bohlen S. Bellevue Hospital Center 2013 Community Health Needs Assessment and Implementation Strategy. New York: New York City Health and Hospitals Corporation URL: http://www.nyc.gov/html/hhc/downloads/pdf/community-assessment/hhc-chna-bellevue.pdf [accessed 2014-12-31] [WebCite Cache ID 6VEe7wB2A]
29. Bradley C, Plowright R, Stewart J, Valentine J, Witthaus E. The Diabetes Treatment Satisfaction Questionnaire change version (DTSQc) evaluated in insulin glargine trials shows greater responsiveness to improvements than the original DTSQ. Health Qual Life Outcomes 2007 Oct;5:57 [FREE Full text] [doi: 10.1186/1477-7525-5-57] [Medline: 17927832]
30. Eysenbach G, Consort- E. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

APCC: Adult Primary Care Center
HHC: Health and Hospitals Corporation
HbA1c: glycated hemoglobin
MITI: Mobile Insulin Titration Intervention
RA: research assistant
SES: socioeconomic status
SMS: short message service

Edited by G Eysenbach; submitted 18.05.15; peer-reviewed by R Dobson; comments to author 13.06.15; revised version received 18.06.15; accepted 29.06.15; published 17.07.15.

Please cite as:
Levy N, Moynihan V, Nilo A, Singer K, Bernik LS, Etiebet MA, Fang Y, Cho J, Natarajan S
The Mobile Insulin Titration Intervention (MITI) for Insulin Adjustment in an Urban, Low-Income Population: Randomized Controlled Trial
J Med Internet Res 2015;17(7):e180
URL: http://www.jmir.org/2015/7/e180/
doi:10.2196/jmir.4716
PMID:26187303

©Natalie Levy, Victoria Moynihan, Annielyn Nilo, Karyn Singer, Lidia S Bernik, Mary-Ann Etiebet, Yixin Fang, James Cho, Sundar Natarajan. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 17.07.2015. 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

Evaluation of Internet-Based Interventions on Waist Circumference Reduction: A Meta-Analysis

Dong-Chul Seo[1], PhD; Jingjing Niu[2], MS

[1]College of Health Sciences, Department of Health Education and Management, Ewha Womans University, Seoul, Republic Of Korea

[2]School of Public Health, Department of Applied Health Science, Indiana University, Bloomington, IN, United States

Corresponding Author:
Dong-Chul Seo, PhD
College of Health Sciences
Department of Health Education and Management
Ewha Womans University
Human Ecology Building 204
52, Ewhayeodae-gil, Seodaemun-gu
Seoul, 120-750
Republic Of Korea
Phone: 82 2 3277 2580
Fax: 82 2 3277 2867
Email: dseo@ewha.ac.kr

Abstract

Background: Internet-based interventions are more cost-effective than conventional interventions and can provide immediate, easy-to-access, and individually tailored support for behavior change. Waist circumference is a strong predictor of an increased risk for a host of diseases, such as hypertension, diabetes, and dyslipidemia, independent of body mass index. To date, no study has examined the effect of Internet-based lifestyle interventions on waist circumference change.

Objective: This study aimed to systematically review the effect of Internet-based interventions on waist circumference change among adults.

Methods: This meta-analysis reviewed randomized controlled trials (N=31 trials and 8442 participants) that used the Internet as a main intervention approach and reported changes in waist circumference.

Results: Internet-based interventions showed a significant reduction in waist circumference (mean change –2.99 cm, 95% CI −3.68 to −2.30, I^2=93.3%) and significantly better effects on waist circumference loss (mean loss 2.38 cm, 95% CI 1.61-3.25, I^2=97.2%) than minimal interventions such as information-only groups. Meta-regression results showed that baseline waist circumference, gender, and the presence of social support in the intervention were significantly associated with waist circumference reduction.

Conclusions: Internet-based interventions have a significant and promising effect on waist circumference change. Incorporating social support into an Internet-based intervention appears to be useful in reducing waist circumference. Considerable heterogeneity exists among the effects of Internet-based interventions. The design of an intervention may have a significant impact on the effectiveness of the intervention.

(J Med Internet Res 2015;17(7):e181) doi:10.2196/jmir.3921

KEYWORDS

waist circumference; obesity; adiposity; Internet; intervention studies

Introduction

The prevalence of obesity has been increasing worldwide for approximately 50 years and has now become a global pandemic [1]. Lifestyle interventions balancing energy intake and energy expenditure have been suggested as effective tools to treat obesity and prevent obesity-related health burdens [2,3]. Internet-based interventions can provide immediate, easy-to-access, and individually tailored support for behavior change, which attract a large number of individuals, including the young or elderly, healthy, disabled or sick, and various ethnicities [4,5]. It has been reported that Internet-based lifestyle

XSL•FO
RenderX

interventions can be as effective as phone- or person-based interventions in reducing body weight [6-8]. In contrast, Internet-based interventions are more cost-effective than conventional interventions [9-12].

Waist circumference, as a simple and effective measure of central obesity, is a strong predictor of an increased risk for hypertension, diabetes mellitus, dyslipidemia, metabolic syndrome, and coronary heart disease independent of body mass index (BMI) [13,14]. Changes in waist circumference in response to lifestyle interventions reflect changes in central obesity [15,16]. Studies have reported that waist circumference can be reduced while no significant changes in body weight occur [17-19]. Few studies, however, have systemically evaluated the effect of lifestyle interventions on waist circumference change. Therefore, this study examined the effect of Internet-based lifestyle interventions on waist circumference change.

Previous reviews have reported that Internet-based interventions can promote physical activity and significantly reduce body weight [2,20,21]. Khaylis et al [22] conducted a systematic review of efficacious technology-based weight-loss interventions and identified self-monitoring, counselor feedback and communication, social support, structured programs, and individually tailored programs as a key to successful interventions. In addition, the literature identified goal setting, motivational interviewing, and incentives as potential factors that increase intervention effectiveness [9,11,23-26]. Seo and Sa [27] also reported that the number of components was associated with the effect of lifestyle interventions. Based on the existing evidence, we hypothesize the following: (1) an Internet-based intervention can significantly reduce waist circumference; (2) Internet-based interventions reduce waist circumference more than conventional minimal interventions, such as those with usual care or information-only delivery; and (3) the number and type of components in lifestyle interventions are significantly associated with the effect on waist circumference change.

Methods

Search Strategy

An electronic search was performed in the following databases: Academic Search Premier, CINAHL Plus with Full Text, Educational Resource Information Center (ERIC), Health Source Nursing/Academic Edition, MEDLINE, PsycARTICLES, SPORTDiscus with Full Text Results, and ProQuest Dissertations and Theses A&I database. The search terms used various combinations of the following keywords or phrases: adiposity, weight, overweight, obese, obesity, lifestyle, nutrition, diet, intake, physical activity, exercise, eHealth, Web, online, email, electronic mail, Internet, social networking, treatment, therapy, interventions, management, trial, waist, central adiposity, random, control, and randomized controlled trial (RCT). After excluding ineligible studies, a manual search was conducted by screening the references of the remaining articles and contacting experts. The detailed search strategy can be found in Multimedia Appendix 1.

Inclusion Criteria

Studies were selected if they met all of the following criteria: (1) published in English peer-reviewed journals between 1980 and April 2014 or dissertations/theses written in English that reported relevant yet unpublished results and were uploaded before April 2014, (2) studies based on RCTs, (3) studies that used the Internet as a major intervention tool in at least 1 arm, (4) studies that used lifestyle interventions (which promote healthy diet, physical activity, or both), (5) studies that reported the mean and standard deviation (SD) or standard error (SE) of the waist circumference, and (6) studies involving adults (aged ≥18 years). Studies were excluded if special diets or medications were used in the intervention or only follow-up data of an intervention were reported.

Data Extraction

The following data were extracted from each included study and substudy: (1) general information, such as the name of the first author and year of publication/completion; (2) characteristics of the substudy, such as intervention location, number of participants, intervention length, frequency, retention rate, participants' compliance, features of the intervention arm, approaches used in adjunction to the Internet (eg, personal contacts via phone, in-person visits, or other devices), intervention content, and whether or not theory, tailoring, self-monitoring and feedback on performance, goal setting, motivational interviewing, social support/social change, and incentives for weight loss were used in the intervention; (3) characteristics of the participants, such as general obesity status, reported existing diseases, mean age, and percentage of male participants; and (4) the mean and SD or SE of the waist circumference at baseline and immediately after the intervention, and the waist circumference change. The SE of the waist circumference change was calculated using the baseline and follow-up SD or SE, assuming an intracorrelation coefficient of 0.5 between pretest and posttest [28], when the SD or SE of the waist circumference change was not reported. The intention-to-treat analysis results were extracted and used when available. The risk of bias was assessed using the Cochrane Collaboration tool [29] and this assessment was used to guide the interpretation of study results.

Data Analysis

Each reported arm was treated as an independent substudy. Treatments that were unlikely to have effects on waist circumference change, such as no intervention, delayed intervention, usual care, and information-only groups, were categorized as "minimal interventions." Paper-, phone-, and person-based interventions were grouped together as "other interventions" because only 6 trials used any of these interventions. We calculated the overall effect sizes of waist circumference changes in Internet-based, minimal, and other interventions. Next, we compared the effects between Internet-based interventions with minimal or other interventions. To examine the effect of a "unique intervention component" on waist circumference changes, intervention components were coded as 0 for the component delivered in both conditions, 1 for the unique component in Internet-based intervention, and 2 for the unique component in minimal intervention. The number

XSL•FO
RenderX

of times each intervention component was uniquely found only in Internet-based interventions was computed.

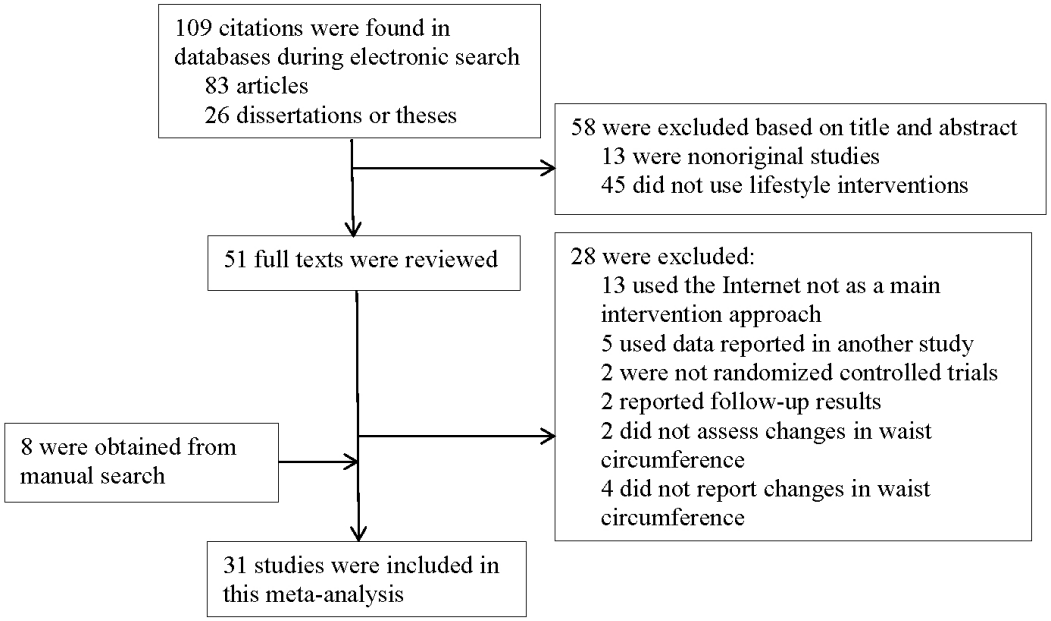
Effect sizes were presented as the mean waist circumference change in centimeters with a 95% confidence interval (CI). Funnel plot and Begg's test were used to test publication bias. The I^2 index was used to test between-study heterogeneity. A meta-regression was performed to identify characteristics that were significantly associated with differences in waist circumference changes between Internet-based interventions and minimal interventions, although it was likely underpowered. Due, in part, to the concern about the possible insignificant findings arising from low power, another meta-regression was performed for the waist circumference changes from baseline to posttest. The adjusted R^2 was calculated to present the predictive power of meta-regression models. Random effects models were used if significant between-study heterogeneity was detected. Analyses were performed using Stata 13 (StataCorp LP, College Station, TX, USA).

Results

Overview

After removing duplicates, the electronic search retrieved 83 articles and 26 dissertations or theses. The manual search retrieved 8 additional studies. Four studies indicated measurement of waist circumference, yet failed to report adequate information on the waist circumference. We contacted the authors but could not obtain additional information necessary for meta-analysis. Thus, these 4 studies were excluded from analysis. Figure 1 demonstrates the flow of the literature search. The studies were reviewed independently by 2 reviewers and any disagreement encountered was resolved by discussion.

Figure 1. Flow chart of literature search.



Characteristics of Included Studies

This review includes 31 intervention trials involving 72 intervention arms and 8442 adults [6-12,23,25,30-51]. The number of total participants ranged from 21 to 1692, and the mean sample size was 272 per study. Four studies only recruited women and 5 studies only recruited men. Fifteen of 31 interventions were conducted in the United States. Among substudy participants, the mean age ranged from 19.0 to 64.9 years and the mean baseline waist circumference ranged from 81.7 to 128.4 cm. The intervention length ranged from 4 weeks to 2 years. The retention rate ranged from 21.7% to 100%, and the mean retention was 75.2%. Characteristics of the included substudies are provided in Table 1.

XSL•FO
RenderX

Table 1. Descriptive data of substudies included in this meta-analysis (N=72).

Author	Year	Arm	Features	N	Waist circumference (cm)		
					Base-line	Post	Δ Mean (SD)
Bennett [9]	2010	Minimal[a]	Brochure	50	N/A	N/A	–1.9 (10.8)
		Internet	Online program + forum	51	N/A	N/A	–1.9 (10.8)
Bischoff [30]	2010	Basic Internet-based	Online program + email contacts	22	94.7	92.6	–2.1 (3.5)
		Enhanced Internet-based	Basic + goal setting	21	83.8	82.5	–1.3 (1.9)
Booth [10]	2008	Basic Internet-based	Online PA program + forum + email feedbacks	26	96.9	N/A	–4.5 (4.5)
		Enhanced Internet-based	Basic + nutrition component	27	95.6	N/A	–3.2 (2.9)
Bukhari [31]	2009	Minimal[a]	One class + a counseling	11	98.8	100.6	1.8 (16.9)
		Internet	Online programs	16	107.5	101.8	–5.7 (14.6)
Carr [32]	2008	Minimal[a]	None	18	99.2	99.8	0.6 (2.1)
		Internet	Online sessions + email contacts	14	100.6	96.6	–4.0 (2.5)
Chambliss [33]	2011	Minimal[a]	None	28	100.1	N/A	–0.6 (5.2)
		Basic Internet-based	Online monitor + email counseling	33	97.1	N/A	–3.4 (4.6)
		Enhanced Internet-based	Basic + behaviorally tailored	34	98.2	N/A	–2.8 (5.4)
Chen [34]	2013	Minimal[a]	None	31	88.9	88.3	–0.6 (10.2)
		Internet	Online program + feedbacks	32	91.9	88.4	–3.5 (11.1)
Chung [35]	2014	Minimal[a]	None	19	94.5	92.6	–1.9 (8.3)
		Paper	Logbook	16	93.2	89.1	–4.1 (7.1)
		Internet	Online logs + evaluation	19	91.9	88.5	–3.4 (11.0)
Collins [36]	2012	Minimal[a]	None	104	107.2	N/A	0.3 (3.1)
		Basic Internet-based	Online programs + forums + email contacts	99	106.9	N/A	–2.6 (4.0)
		Enhanced Internet-based	Basic + personalized + feedbacks	106	106.6	N/A	–3.2 (5.0)
Dekkers [37]	2011	Minimal[a]	None	49	101.7	99.2	–2.5 (8.9)
		Phone	Phone sessions + counseling	44	99.9	96.4	–3.5 (10.6)
		Internet	Online sessions + email counseling	48	102.9	99.4	–3.5 (11.3)
Hansen [38]	2012	Minimal[a]	None	585	89.6	89.1	–0.5 (8.4)
		Internet	Online program + forum	583	90.1	90.0	–0.1 (8.5)
Herrick [39]	2009	Minimal	None	860	81.7	N/A	0.7 (4.7)
		Internet	Online programs + email reminders	832	81.9	N/A	0.3 (2.9)
Hunter [40]	2008	Minimal[a]	None	222	94.2	93.4	–0.4 (3.8)
		Internet	Online program + feedbacks	224	94.5	92.2	–2.1 (4.3)
Kang [41]	2010	Minimal[a]	General information	75	85.4	88.6	3.2 (8.8)
		Basic Internet-based	1-year email education	25	83.2	84.3	1.1 (5.5)
		Enhanced Internet-based	2-year email education	25	89.1	87.3	–1.8 (3.0)
Mehring [23]	2013	Minimal[a]	Usual care	77	110.9	104.4	–6.9 (6.9)
		Internet	Online programs	109	107.3	106.6	–2.4 (5.0)
Mobley [42]	2006	Basic in-person	In-person counseling	32	100.7	99.3	–1.4 (9.1)
		Enhanced in-person	In-person counseling	33	101.8	98.5	–3.3 (9.7)

XSL•FO
RenderX

Author	Year	Arm	Features	N	Waist circumference (cm)		
					Baseline	Post	Δ Mean (SD)
		Basic Internet-based	Online programs	29	102.7	102.4	–0.3 (10.1)
		Enhanced Internet-based	Online programs	29	101.6	103.4	1.8 (8.4)
Morgan [11]	2009	Minimal[a]	None	31	102.8	N/A	–5.2 (5.4)
		Internet	Online programs + notice board + email feedbacks	34	103.4	N/A	–4.4 (5.7)
Morgan [6]	2011	Minimal	None	45	99.4	N/A	1.5 (4.5)
		Internet	Online program + email feedbacks	65	101.6	N/A	–4.4 (4.8)
Morgan [43]	2013	Minimal	None	52	113.6	N/A	–0.8 (2.9)
		Paper	Books	54	112.6	N/A	–3.7 (4.5)
		Internet	Online programs + email feedbacks	53	113.7	N/A	–5.4 (5.2)
Patrick [12]	2011	Minimal[a]	Website with general health information	217	112.9	111.6	–1.3 (11.4)
		Internet	Online assessment + sessions + feedbacks + email counseling	224	113.7	112.1	–1.6 (11.4)
Pressler [44]	2010	Basic Internet-based	Nonstructured	27	101.9	98.3	–3.6 (8.6)
		Enhanced Internet-based	Structured	50	100.5	98.0	–2.5 (7.8)
Pullen [45]	2008	Basic Internet-based	Online program	8	98.1	96.2	–1.9 (6.5)
		Enhanced Internet-based	Basic + online discussions	8	91.1	85.7	–5.4 (6.4)
Rogers [7]	2012	Paper	Group sessions	14	125.9	121.9	–4.0 (8.1)
		Basic Internet-based	Online programs	12	128.4	121.6	–6.8 (9.3)
		Enhanced Internet-based	Basic + Bluetooth	13	126.3	120.1	–6.2 (12.6)
Seely [46]	2013	Minimal[a]	None	13	N/A	N/A	–2.8 (2.3)
		Internet	Facebook support group	11	N/A	N/A	–2.9 (3.0)
Tate [47]	2001	Basic Internet-based	Online program + email reminders	45	98.4	N/A	–3.1 (4.4)
		Enhanced Internet-based	Basic + online behavioral therapy	46	98.5	N/A	–6.4 (5.5)
Tate [48]	2003	Basic Internet-based	Online program + email reminders	46	111.0	N/A	–4.4 (5.7)
		Enhanced Internet-based	Basic + email counseling	46	108.0	N/A	–7.2 (7.5)
van Genugten [49]	2012	Basic Internet-based	General information online	239	95.7	93.2	–2.5 (8.6)
		Enhanced Internet-based	Basic + computer tailored + forum	241	95.9	94.4	–1.5 (9.7)
van Wier [8]	2009	Minimal[a]	Brochure	231	101.5	99.5	–2.0 (9.9)
		Internet	Online program + email counseling	236	102.6	98.6	–1.9 (6.3)
		Phone	Phone counseling	235	101.5	98.2	–1.2 (11.7)
Webber [50]	2010	Basic Internet-based	Online program + message board	36	96.5	N/A	–3.6 (5.2)
		Enhanced Internet-based	Basic + motivational treatment	34	97.3	N/A	–3.6 (5.2)
Wijsman [25]	2013	Minimal[a]	None	112	101.4	N/A	–1.3 (3.6)
		Internet	Online program + email counseling	114	102.3	N/A	–2.3 (3.8)
Yoo [51]	2009	Minimal[a]	None	54	91.3	89.1	–2.2 (7.5)
		Internet	Online monitoring + feedbacks	57	89.5	86.8	–2.7 (9.8)

[a] Minimal arm includes control, wait-list, usual care groups or the group that only received standard health information.

Of the 72 intervention arms reviewed in the current study, 33 adapted behavioral theories or therapy principles, 40 prompted self-monitoring of behavior, 39 used feedback on performance and individual tailoring or counseling, 31 used goal setting, 15 planned online social support/social change, 6 used motivational interviewing, and 2 used incentives to encourage weight loss.

XSL•FO
RenderX

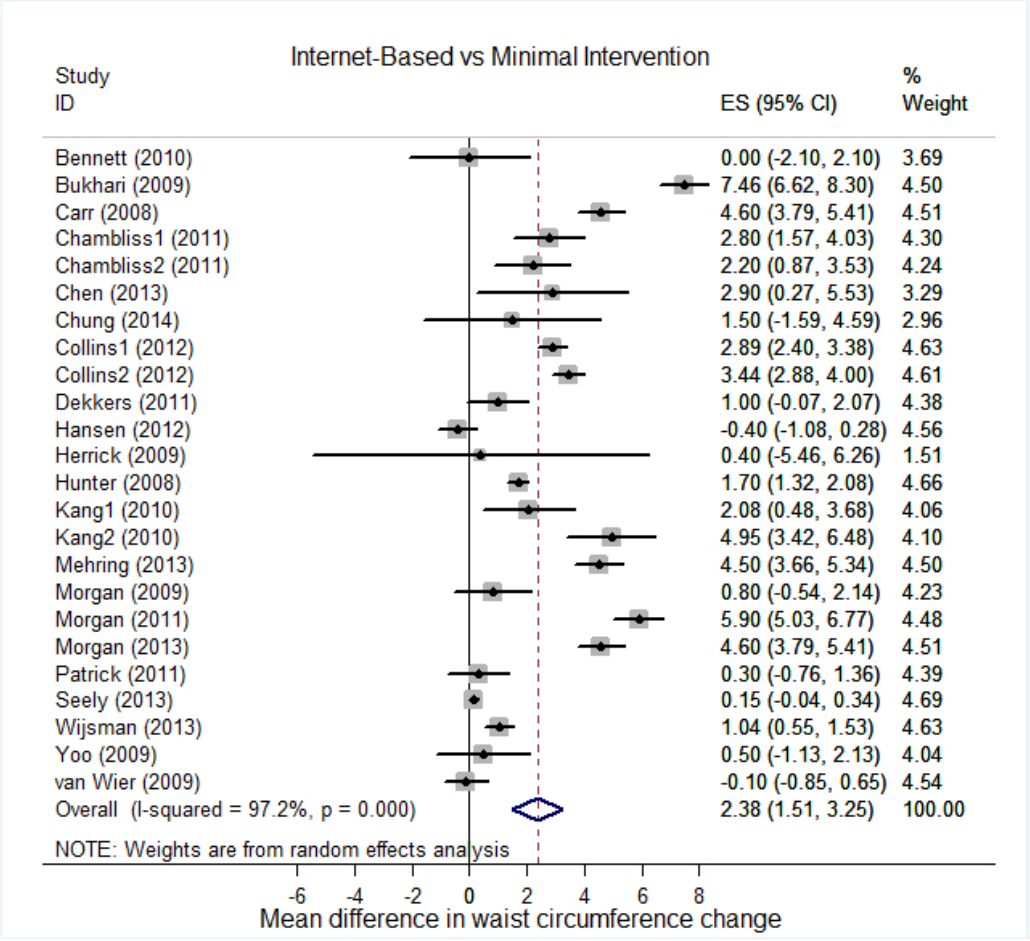
Regarding the total number of components used in each arm, approximately one-third of the arms used none, one-third used 1 to 3 components, and the final one-third used 4 to 6 components. Details about the arm components can be found in Multimedia Appendix 2. A total of 24 pairs were extracted for comparison between Internet-based interventions and minimal interventions. The number of times each component was uniquely found only in Internet-based interventions was 16 for theory, 21 for tailoring, 21 for monitoring, 15 for goal setting, 4 for motivational interviewing, 11 for social support, and 1 for incentive.

Bias assessment showed the following results: 17 studies provided details on random sequence generation, 18 studies provided details on allocation concealment, only 2 studies reported blinding participants, and 9 studies reported blinding assessors. As shown in Multimedia Appendix 3, the bias assessment indicated no evidence of selective reporting of outcomes. As shown in the funnel plot for publication bias (see Multimedia Appendix 4), no significant publication bias was detected (P=.31 for Begg's test) for Internet-based interventions as evaluated by the waist circumference change in each study arm. Multimedia Appendix 5 shows content and supplementary approaches of substudies included in this meta-analysis.

Overall Effects of Interventions

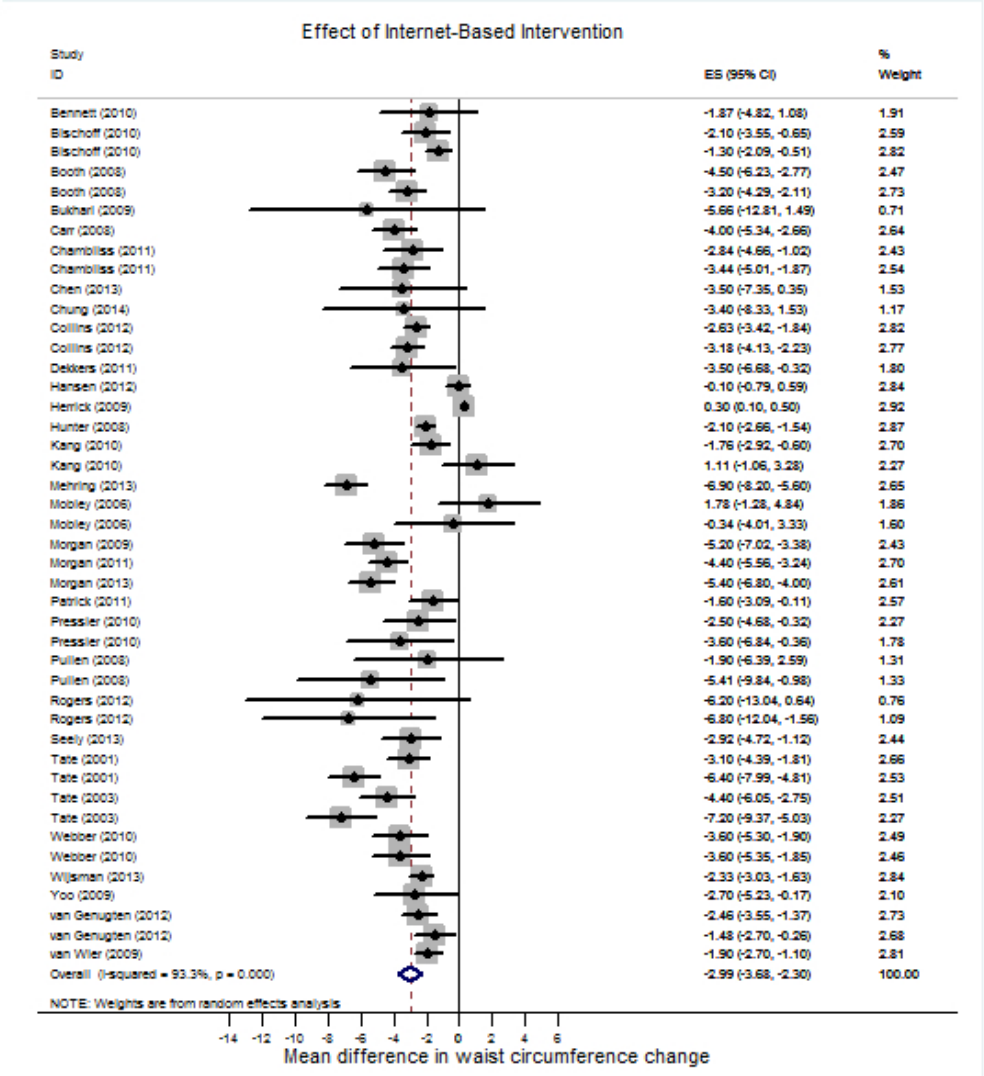
Figure 2 shows the differences in waist circumference change between Internet-based interventions and minimal interventions. Internet-based interventions showed significantly better effects on waist circumference reduction (mean change 2.38 cm, 95% CI 1.51-3.25) compared with minimal interventions. Few differences were observed with respect to the waist circumference change between Internet-based interventions and paper-, phone-, or person-based interventions (mean change −0.61 cm, 95% CI −2.05 to 0.83, P=.42). Figure 3 provides a forest plot representing the effect size of Internet-based interventions on the waist circumference change. Overall, Internet-based interventions significantly reduced the waist circumference (mean change −2.99 cm, 95% CI −3.68 to −2.30), whereas minimal interventions (mean change −0.81 cm, 95% CI −1.41 to −0.20) and other interventions (mean −2.82 cm, 95% CI −3.89 to −1.74) also reduced waist circumference. Large and significant between-study heterogeneity was observed (I^2=93.3%, P<.001).

Figure 2. Forest plot for the differences in waist circumference changes between Internet-based interventions and minimal interventions. % Weight: weights assigned to substudies.



XSL•FO
RenderX

Figure 3. Forest plot for the effect of Internet-based intervention on waist circumference changes. % Weight: weights assigned to substudies.



Characteristics Associated With the Waist Circumference Change

The meta-regression of differences in waist circumference changes (changes in minimal intervention groups minus changes in Internet-based intervention groups) showed no significant associations between effect sizes with the content and number of unique intervention components. To further investigate the effect of intervention components, a meta-regression of waist circumference changes from baseline to posttest was conducted. Results of this meta-regression are shown in Table 2. Stepwise meta-regression showed that only the mean waist circumference at baseline (coefficient = −0.16, P<.001) and proportion of male participants (coefficient = −0.02, P=.02) were significantly associated with the effect on waist circumference change (I^2=69.8%, P<.001) among the characteristics of the participants that included mean age, reported existing diseases, and status of general obesity. The waist circumference at baseline alone explained 45.3% of the between-study variation in waist circumference changes. Controlling for the baseline waist circumference and proportion of male participants (R^2=.58), the component of social support in Internet-based intervention was associated with a significantly better effect on waist circumference changes (mean difference −1.16 cm, P=.03), which increased the R^2 to .66.

XSL•FO
RenderX

Table 2. Meta-regression of waist circumference change from baseline to posttest.[a]

Intervention component	Coefficient	SE	*P*	I^2
Mean waist circumference at baseline (cm)	–0.16	0.03	<.001	69.8%
Male participants (%)	0.02	0.01	.02	
Length (weeks)[b]	0.01	0.01	.89	69.9%
Frequency (week per contact)[b]	–0.13	0.20	.54	71.4%
Retention (%)[b]	–0.02	0.01	.15	71.1%
Intervention content[b]				70.6%
Physical activity vs both	0.34	0.58	.56	
Nutrition vs both	–1.38	2.27	.55	
Theory (yes vs no)[b]	0.38	0.53	.48	69.7%
Tailoring (yes vs no)[b]	–0.45	0.58	.45	69.9%
Monitoring (yes vs no)[b]	–0.62	0.66	.35	69.6%
Goal setting (yes vs no) [b]	0.50	0.57	.38	69.6%
Motivational interviewing (yes vs no) [b]	–0.09	0.77	.91	69.5%
Social support (yes vs no)[b]	–1.16	0.51	.03	66.9%
Number of components[b]	–0.12	0.16	.48	69.6%

[a] The incentive variable was not included in the model because only 1 study used incentives to encourage weight loss.

[b] Controlled for mean waist circumference at baseline and percentage of male participants.

Discussion

This study was the first attempt to the authors' knowledge that evaluated effect sizes of Internet-based lifestyle interventions on decreasing waist circumference. This meta-analysis showed that Internet-based interventions not only decreased waist circumference substantially at posttest (a mean decrease of 2.99 cm), but also did so significantly more than minimal interventions. Given that a meta-analysis of workplace physical activity and dietary behavioral interventions only demonstrated an average waist circumference reduction of 0.67 cm (95% CI −1.96 to 0.63) [52] and another meta-analysis of antiobesity drugs showed an additional waist circumference reduction of 1.72 to 3.58 cm at 3 months among overweight or obese adults compared with standard care groups [53], Internet-based interventions appear to have a promising effect on waist circumference reduction. It deserves mention that this meta-analysis found few differences in waist circumference reduction between Internet-based intervention and paper-, phone-, or person-based lifestyle interventions. Considering that Internet-based interventions can attract a larger number of individuals with various backgrounds [2,4,54], can provide immediate and easy-access support at a lower cost [21,55], are more accessible to older adults and residents of geographically isolated communities [6,11,26], and are less obtrusive [26] than traditional methods, the substantial effect of Internet-based lifestyle interventions on the waist circumference change found in this study adds to the reason that Internet-based rather than traditional lifestyle interventions should be more widely and boldly explored.

Of the 31 trials reviewed in this study, 24 differences between Internet-based intervention and minimal intervention were identified. Compared with minimal interventions, Internet-based interventions included 1 to 6 unique intervention components. In addition, 13 trials tested the differences between basic and enhanced Internet-based interventions. Compared with the basic interventions, the enhanced intervention included 1 or 2 additional intervention features, such as adding healthy diet promotion to physical activity promotion or adding Bluetooth technology to the basic intervention. As indicated in Table 1, some enhanced interventions had a better effect on waist circumference change whereas others had a worse effect than basic interventions. The result of this study indicates that no conclusive evaluation is warranted on the efficacy of such additional features of the Internet-based lifestyle interventions.

To complement findings drawn from current and previous systematic reviews, we tested the associations between key intervention characteristics and waist circumference reduction in meta-regression models where independent variables were selected based on previous findings [9,12,22,27]. We examined whether the content and number of unique intervention components could adequately predict the differences in waist circumference changes between Internet-based interventions and minimal interventions only to find no significant association. The lack of significant association may be explained by lack of power due to the small sample size (n=24). It is also possible that the effect size can be explained by intervention features not tested in the current study. Similar future meta-analysis research is warranted that includes more studies and possibly a different framework. Due, in part, to the concern about the possible

XSL•FO
RenderX

insignificant findings arising from low power, another meta-regression was performed for the waist circumference changes from baseline to posttest. We found that only the availability of social support was significantly associated with the waist circumference change after controlling for the main characteristics of participants. This means that providing sufficient social support is important to improve the efficacy of Internet-based lifestyle interventions. The lack of significant associations between waist circumference reduction and intervention length, intervention topic (nutrition only, physical activity only, or both), and the approach used in adjunction to the Internet (eg, personal contacts via phone, in-person visits, or use of such devices as Bluetooth, pedometer, or accelerometer) deserves further research. Although other intervention characteristics did not yield significant results in this review, further investigations are needed to draw conclusive suggestions.

It is worth noting that considerable heterogeneity remained after controlling for baseline waist circumference, gender, and intervention components identified by this study. It indicates that there is heterogeneity in effect sizes among the Internet-based interventions examined in this review that has yet to be accounted for. This may have to do with lack of frameworks that informed the design of Internet-based interventions reviewed in this study or lack of use of well-defined constructs or concepts. Previous studies have found that there is a lack of framework for the design of technology-based behavioral interventions and each research team used their own ways to develop and report technology-based interventions [56,57]. As a result, many of such technology-based intervention features lack comparability between different studies. Eysenbach and colleagues [58] developed the CONSORT-EHEALTH to standardize reports of eHealth/mHealth interventions, which has been very helpful in disseminating and comparing research reports. Recently, Schueller et al [56] proposed the modular system Purple to assist the development of Internet-based and mobile-based applications for health behavior change and Mohr et al [57] proposed a comprehensive framework, the Behavioral Intervention Technology (BIT) Model. These recently proposed frameworks and models should be fully utilized to inform the design of future technology-based interventions. In addition, for comparability and clarity of findings in behavior change interventions, it is desirable to use well-defined terms such as those shown in the Coventry, Aberdeen & London-Refined (CALO-RE) taxonomy [59]. These new frameworks and taxonomy of behavior change techniques will help increase comparability between different technology-based behavioral intervention studies as well as enhance the effectiveness of such interventions.

This meta-analysis review has the following limitations. First, gender-specific analyses were not performed due to a lack of gender-specific information, although baseline waist circumference and changes in the waist circumference may differ by gender. Future studies would be desirable that investigate gender-specific waist circumference changes of Internet-based interventions. Second, paper-, phone-, and person-based interventions were grouped together in this analysis due to a lack of data. Future research can be conducted to compare each mode of lifestyle interventions with other modes of interventions in terms of effect size when the sample size is appropriate. Third, the effect of compliance and incentives were not investigated in this study due to a lack of such information in the reviewed studies. Researchers are recommended to report information on participant compliance and incentives. Finally, considerable heterogeneity in waist circumference changes remained after controlling for covariates including baseline waist circumference and gender. This might indicate that reviewed interventions lacked frameworks that informed their study design. Thus, it is possible that not all the efficacious intervention components in reducing waist circumference for Internet-based lifestyle interventions might have been examined and analyzed.

In summary, Internet-based lifestyle interventions showed a significant and substantial effect on waist circumference change. Internet-based interventions showed comparable effects on the waist circumference change to paper-, phone-, and person-based interventions. Online social support appears to strengthen the effect of Internet-based programs on waist circumference reduction. Internet-based programs are recommended for obesity or lifestyle as effective and efficient interventions. It is also recommended to integrate online social support into Internet-based programs to achieve better effects on weight control.

Acknowledgments

This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2012S1A3A2033416). The funding agency was not involved with conceptualization, study design, analysis, or drafting of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategy.

[PDF File (Adobe PDF File), 17KB - jmir_v17i7e181_app1.pdf]

XSL•FO
RenderX

Multimedia Appendix 2

Components of sub-studies included in this meta-analysis.

[PDF File (Adobe PDF File), 130KB - jmir_v17i7e181_app2.pdf]

Multimedia Appendix 3

Table for the risk of bias assessment.

[PDF File (Adobe PDF File), 25KB - jmir_v17i7e181_app3.pdf]

Multimedia Appendix 4

Funnel plot for publication bias.

[PDF File (Adobe PDF File), 41KB - jmir_v17i7e181_app4.pdf]

Multimedia Appendix 5

Content and supplementary approaches of substudies included in this meta-analysis.

[PDF File (Adobe PDF File), 29KB - jmir_v17i7e181_app5.pdf]

References

1. Swinburn BA, Sacks G, Hall KD, McPherson K, Finegood DT, Moodie ML, et al. The global obesity pandemic: shaped by global drivers and local environments. Lancet 2011 Aug 27;378(9793):804-814. [doi: 10.1016/S0140-6736(11)60813-1] [Medline: 21872749]
2. Kodama S, Saito K, Tanaka S, Horikawa C, Fujiwara K, Hirasawa R, et al. Effect of Web-based lifestyle modification on weight control: a meta-analysis. Int J Obes (Lond) 2012 May;36(5):675-685. [doi: 10.1038/ijo.2011.121] [Medline: 21694698]
3. Seo DC, Sa J. A meta-analysis of psycho-behavioral obesity interventions among US multiethnic and minority adults. Prev Med 2008 Dec;47(6):573-582. [doi: 10.1016/j.ypmed.2007.12.010] [Medline: 18201758]
4. 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]
5. Norman GJ, Zabinski MF, Adams MA, Rosenberg DE, Yaroch AL, Atienza AA. A review of eHealth interventions for physical activity and dietary behavior change. Am J Prev Med 2007 Oct;33(4):336-345 [FREE Full text] [doi: 10.1016/j.amepre.2007.05.007] [Medline: 17888860]
6. Morgan PJ, Collins CE, Plotnikoff RC, Cook AT, Berthon B, Mitchell S, et al. Efficacy of a workplace-based weight loss program for overweight male shift workers: the Workplace POWER (Preventing Obesity Without Eating like a Rabbit) randomized controlled trial. Prev Med 2011 May;52(5):317-325. [doi: 10.1016/j.ypmed.2011.01.031] [Medline: 21300083]
7. Rogers R. The Comparison of a Technology-Based System and In-Person Behavioral Weight Loss Intervention in the Severely Obese [dissertation]. Pittsburg, PA: University of Pittsburgh; 2012.
8. van Wier MF, Ariëns GA, Dekkers JC, Hendriksen IJ, Smid T, van Mechelen W. Phone and e-mail counselling are effective for weight management in an overweight working population: a randomized controlled trial. BMC Public Health 2009;9:6 [FREE Full text] [doi: 10.1186/1471-2458-9-6] [Medline: 19134171]
9. Bennett GG, Herring SJ, Puleo E, Stein EK, Emmons KM, Gillman MW. Web-based weight loss in primary care: a randomized controlled trial. Obesity (Silver Spring) 2010 Feb;18(2):308-313 [FREE Full text] [doi: 10.1038/oby.2009.242] [Medline: 19696764]
10. Booth AO, Nowson CA, Matters H. Evaluation of an interactive, Internet-based weight loss program: a pilot study. Health Educ Res 2008 Jun;23(3):371-381 [FREE Full text] [doi: 10.1093/her/cyn007] [Medline: 18349032]
11. Morgan PJ, Lubans DR, Collins CE, Warren JM, Callister R. The SHED-IT randomized controlled trial: evaluation of an Internet-based weight-loss program for men. Obesity (Silver Spring) 2009 Nov;17(11):2025-2032. [doi: 10.1038/oby.2009.85] [Medline: 19343018]
12. Patrick K, Calfas KJ, Norman GJ, Rosenberg D, Zabinski MF, Sallis JF, et al. Outcomes of a 12-month web-based intervention for overweight and obese men. Ann Behav Med 2011 Dec;42(3):391-401. [doi: 10.1007/s12160-011-9296-7] [Medline: 21822750]
13. Berglund G, Ljungman S, Hartford M, Wilhelmsen L, Björntorp P. Type of obesity and blood pressure. Hypertension 1982;4(5):692-696 [FREE Full text] [Medline: 7106936]
14. Wang Y, Rimm EB, Stampfer MJ, Willett WC, Hu FB. Comparison of abdominal adiposity and overall obesity in predicting risk of type 2 diabetes among men. Am J Clin Nutr 2005 Mar;81(3):555-563 [FREE Full text] [Medline: 15755822]

XSL•FO
RenderX

15. Kamel EG, McNeill G, Van Wijk MC. Change in intra-abdominal adipose tissue volume during weight loss in obese men and women: correlation between magnetic resonance imaging and anthropometric measurements. Int J Obes Relat Metab Disord 2000 May;24(5):607-613. [Medline: 10849583]
16. Ross R, Rissanen J, Hudson R. Sensitivity associated with the identification of visceral adipose tissue levels using waist circumference in men and women: effects of weight loss. Int J Obes Relat Metab Disord 1996 Jun;20(6):533-538. [Medline: 8782729]
17. Ross R, Janssen I, Dawson J, Kungl A, Kuk JL, Wong SL, et al. Exercise-induced reduction in obesity and insulin resistance in women: a randomized controlled trial. Obes Res 2004 May;12(5):789-798. [doi: 10.1038/oby.2004.95] [Medline: 15166299]
18. Lee S, Kuk JL, Davidson LE, Hudson R, Kilpatrick K, Graham TE, et al. Exercise without weight loss is an effective strategy for obesity reduction in obese individuals with and without Type 2 diabetes. J Appl Physiol (1985) 2005 Sep;99(3):1220-1225 [FREE Full text] [doi: 10.1152/japplphysiol.00053.2005] [Medline: 15860689]
19. Gan SK, Kriketos AD, Ellis BA, Thompson CH, Kraegen EW, Chisholm DJ. Changes in aerobic capacity and visceral fat but not myocyte lipid levels predict increased insulin action after exercise in overweight and obese men. Diabetes Care 2003 Jun;26(6):1706-1713. [Medline: 12766098]
20. Vandelanotte C, Spathonis KM, Eakin EG, Owen N. Website-delivered physical activity interventions a review of the literature. Am J Prev Med 2007 Jul;33(1):54-64. [doi: 10.1016/j.amepre.2007.02.041] [Medline: 17572313]
21. van den Berg MH, Schoones JW, Vliet Vlieland TP. Internet-based physical activity interventions: a systematic review of the literature. J Med Internet Res 2007;9(3):e26 [FREE Full text] [doi: 10.2196/jmir.9.3.e26] [Medline: 17942388]
22. Khaylis A, Yiaslas T, Bergstrom J, Gore-Felton C. A review of efficacious technology-based weight-loss interventions: five key components. Telemed J E Health 2010 Nov;16(9):931-938 [FREE Full text] [doi: 10.1089/tmj.2010.0065] [Medline: 21091286]
23. Mehring M, Haag M, Linde K, Wagenpfeil S, Frensch F, Blome J, et al. Effects of a general practice guided web-based weight reduction program--results of a cluster-randomized controlled trial. BMC Fam Pract 2013;14:76 [FREE Full text] [doi: 10.1186/1471-2296-14-76] [Medline: 23981507]
24. Miller W, Rollnick S. Motivational Interviewing, Third Edition: Helping People Change. New York: The Guilford Press; 2012.
25. 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]
26. Wing RR, Jeffery RW. Benefits of recruiting participants with friends and increasing social support for weight loss and maintenance. J Consult Clin Psychol 1999 Feb;67(1):132-138. [Medline: 10028217]
27. Seo DC, Sa J. A meta-analysis of obesity interventions among U.S. minority children. J Adolesc Health 2010 Apr;46(4):309-323. [doi: 10.1016/j.jadohealth.2009.11.202] [Medline: 20307819]
28. Follmann D, Elliott P, Suh I, Cutler J. Variance imputation for overviews of clinical trials with continuous response. J Clin Epidemiol 1992 Jul;45(7):769-773. [Medline: 1619456]
29. Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928 [FREE Full text] [Medline: 22008217]
30. Bischoff S. The Effects of Goal Setting on Increasing Physical Activity Behavior Through a 16-week Internet-delivered Intervention [master's thesis]. Laramie, WY: University of Wyoming; 2010.
31. Chambliss HO, Huber RC, Finley CE, McDoniel SO, Kitzman-Ulrich H, Wilkinson WJ. Computerized self-monitoring and technology-assisted feedback for weight loss with and without an enhanced behavioral component. Patient Educ Couns 2011 Dec;85(3):375-382. [doi: 10.1016/j.pec.2010.12.024] [Medline: 21295433]
32. Bukhari A. Effectiveness of a Web-based Nutrition Education Program to Reduce Cardiovascular Disease Risk Among US Army Personnel and Their Families ("Defend Your Heart" study) [dissertation]. Fort Collins, CO: Colorado State University; 2009.
33. Carr LJ, Bartee RT, Dorozynski C, Broomfield JF, Smith ML, Smith DT. Internet-delivered behavior change program increases physical activity and improves cardiometabolic disease risk factors in sedentary adults: results of a randomized controlled trial. Prev Med 2008 May;46(5):431-438. [doi: 10.1016/j.ypmed.2007.12.005] [Medline: 18207228]
34. Chen YC, Tsao LI, Huang CH, Yu YY, Liu IL, Jou HJ. An Internet-based health management platform may effectively reduce the risk factors of metabolic syndrome among career women. Taiwan J Obstet Gynecol 2013 Jun;52(2):215-221. [doi: 10.1016/j.tjog.2013.04.011] [Medline: 23915854]
35. Collins CE, Morgan PJ, Jones P, Fletcher K, Martin J, Aguiar EJ, et al. A 12-week commercial web-based weight-loss program for overweight and obese adults: randomized controlled trial comparing basic versus enhanced features. J Med Internet Res 2012;14(2):e57 [FREE Full text] [doi: 10.2196/jmir.1980] [Medline: 22555246]
36. Chung LM, Law QP, Fong SS, Chung JW. Teledietetics improves weight reduction by modifying eating behavior: a randomized controlled trial. Telemed J E Health 2014 Jan;20(1):55-62 [FREE Full text] [doi: 10.1089/tmj.2013.0104] [Medline: 24205807]

XSL•FO
RenderX

37. Dekkers JC, van Wier MF, Ariëns GA, Hendriksen IJ, Pronk NP, Smid T, et al. Comparative effectiveness of lifestyle interventions on cardiovascular risk factors among a Dutch overweight working population: a randomized controlled trial. BMC Public Health 2011;11(1):49 [FREE Full text] [doi: 10.1186/1471-2458-11-49] [Medline: 21261935]
38. Hansen AW, Grønbæk M, Helge JW, Severin M, Curtis T, Tolstrup JS. Effect of a Web-based intervention to promote physical activity and improve health among physically inactive adults: a population-based randomized controlled trial. J Med Internet Res 2012;14(5):e145 [FREE Full text] [doi: 10.2196/jmir.2109] [Medline: 23111127]
39. Herrick M. Characterization of Time and Order of Internet Lessons in Project WebHealth [master's thesis]. Brookings, SD: South Dakota State University; 2009.
40. Hunter CM, Peterson AL, Alvarez LM, Poston WC, Brundige AR, Haddock CK, et al. Weight management using the internet a randomized controlled trial. Am J Prev Med 2008 Feb;34(2):119-126. [doi: 10.1016/j.amepre.2007.09.026] [Medline: 18201641]
41. Kang JY, Cho SW, Sung SH, Park YK, Paek YM, Choi TI. Effect of a continuous diabetes lifestyle intervention program on male workers in Korea. Diabetes Res Clin Pract 2010 Oct;90(1):26-33. [doi: 10.1016/j.diabres.2010.06.006] [Medline: 20621377]
42. Mobley A. Evaluation of Behavioral Theory and Integrated Internet/telephone Technologies to Support Military Obesity and Weight Management Program [dissertation]. College Park, MD: University of Maryland; 2006.
43. Morgan PJ, Callister R, Collins CE, Plotnikoff RC, Young MD, Berry N, et al. The SHED-IT community trial: a randomized controlled trial of internet- and paper-based weight loss programs tailored for overweight and obese men. Ann Behav Med 2013 Apr;45(2):139-152. [doi: 10.1007/s12160-012-9424-z] [Medline: 23129021]
44. Pressler A, Knebel U, Esch S, Kölbl D, Esefeld K, Scherr J, et al. An internet-delivered exercise intervention for workplace health promotion in overweight sedentary employees: a randomized trial. Prev Med 2010;51(3-4):234-239. [doi: 10.1016/j.ypmed.2010.07.008] [Medline: 20638409]
45. Pullen CH, Hageman PA, Boeckner L, Walker SN, Oberdorfer MK. Feasibility of Internet-delivered weight loss interventions among rural women ages 50-69. J Geriatr Phys Ther 2008;31(3):105-112. [Medline: 19856616]
46. Seely E. Facebook as Social Support for a One-time Weight Loss Intervention Among College Students [Master's thesis]. Carbondale, IL: Southern Illinois University at Carbondale; 2013.
47. Tate DF, Jackvony EH, Wing RR. Effects of Internet behavioral counseling on weight loss in adults at risk for type 2 diabetes: a randomized trial. JAMA 2003 Apr 9;289(14):1833-1836. [doi: 10.1001/jama.289.14.1833] [Medline: 12684363]
48. 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]
49. van Genugten L, van Empelen P, Boon B, Borsboom G, Visscher T, Oenema A. Results from an online computer-tailored weight management intervention for overweight adults: randomized controlled trial. J Med Internet Res 2012;14(2):e44 [FREE Full text] [doi: 10.2196/jmir.1901] [Medline: 22417813]
50. Webber KH, Gabriele JM, Tate DF, Dignan MB. The effect of a motivational intervention on weight loss is moderated by level of baseline controlled motivation. Int J Behav Nutr Phys Act 2010;7:4 [FREE Full text] [doi: 10.1186/1479-5868-7-4] [Medline: 20157441]
51. Yoo HJ, Park MS, Kim TN, Yang SJ, Cho GJ, Hwang TG, et al. A Ubiquitous Chronic Disease Care system using cellular phones and the internet. Diabet Med 2009 Jun;26(6):628-635. [doi: 10.1111/j.1464-5491.2009.02732.x] [Medline: 19538239]
52. Verweij LM, Coffeng J, van Mechelen W, Proper KI. Meta-analyses of workplace physical activity and dietary behaviour interventions on weight outcomes. Obes Rev 2011 Jun;12(6):406-429. [doi: 10.1111/j.1467-789X.2010.00765.x] [Medline: 20546142]
53. Chilton M, Dunkley A, Carter P, Davies MJ, Khunti K, Gray LJ. The effect of antiobesity drugs on waist circumference: a mixed treatment comparison. Diabetes Obes Metab 2014 Mar;16(3):237-247. [doi: 10.1111/dom.12198] [Medline: 23964810]
54. Hamel LM, Robbins LB, Wilbur J. Computer- and web-based interventions to increase preadolescent and adolescent physical activity: a systematic review. J Adv Nurs 2011 Feb;67(2):251-268. [doi: 10.1111/j.1365-2648.2010.05493.x] [Medline: 21198800]
55. Hansen AW, Grønbæk M, Helge JW, Severin M, Curtis T, Tolstrup JS. Effect of a Web-based intervention to promote physical activity and improve health among physically inactive adults: a population-based randomized controlled trial. J Med Internet Res 2012;14(5):e145 [FREE Full text] [doi: 10.2196/jmir.2109] [Medline: 23111127]
56. Schueller SM, Begale M, Penedo FJ, Mohr DC. Purple: a modular system for developing and deploying behavioral intervention technologies. J Med Internet Res 2014;16(7):e181 [FREE Full text] [doi: 10.2196/jmir.3376] [Medline: 25079298]
57. Mohr DC, Schueller SM, Montague E, Burns MN, Rashidi P. The behavioral intervention technology model: an integrated conceptual and technological framework for eHealth and mHealth interventions. J Med Internet Res 2014;16(6):e146 [FREE Full text] [doi: 10.2196/jmir.3077] [Medline: 24905070]
58. Eysenbach G. CONSORT-EHEALTH: implementation of a checklist for authors and editors to improve reporting of web-based and mobile randomized controlled trials. Stud Health Technol Inform 2013;192:657-661. [Medline: 23920638]

XSL•FO
RenderX

59. Michie S, Ashford S, Sniehotta FF, Dombrowski SU, Bishop A, French DP. A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: the CALO-RE taxonomy. Psychol Health 2011 Nov;26(11):1479-1498. [doi: 10.1080/08870446.2010.540664] [Medline: 21678185]

Abbreviations

BIT: Behavioral Intervention Technology
ERIC: Educational Resource Information Center
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 06.10.14; peer-reviewed by M Mehring, L van Genugten; comments to author 11.02.15; revised version received 05.03.15; accepted 31.05.15; published 21.07.15.

Please cite as:
Seo DC, Niu J
Evaluation of Internet-Based Interventions on Waist Circumference Reduction: A Meta-Analysis
J Med Internet Res 2015;17(7):e181
URL: http://www.jmir.org/2015/7/e181/
doi:10.2196/jmir.3921
PMID:26199208

©Dong-Chul Seo, Jingjing Niu. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 21.07.2015. 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

It's LiFe! Mobile and Web-Based Monitoring and Feedback Tool Embedded in Primary Care Increases Physical Activity: A Cluster Randomized Controlled Trial

Sanne van der Weegen[1*], MSc; Renée Verwey[1,2*], RN, MSc; Marieke Spreeuwenberg[1], MSc, PhD; Huibert Tange[3], MD, PhD; Trudy van der Weijden[3], MD, PhD; Luc de Witte[1,2], MD, PhD

[1]CAPHRI School for Public Health and Primary Care, Department Health Services Research, Maastricht University, Maastricht, Netherlands

[2]Research Centre Technology in Care, Zuyd University of Applied Sciences, Heerlen, Netherlands

[3]CAPHRI School for Public Health and Primary Care, Department Family Medicine, Maastricht University, Maastricht, Netherlands

[*]these authors contributed equally

Corresponding Author:
Sanne van der Weegen, MSc
CAPHRI School for Public Health and Primary Care
Department Health Services Research
Maastricht University
PO Box 616
Maastricht, 6200 MD
Netherlands
Phone: 31 642756346
Fax: 31 433884162
Email: s.vanderweegen@vilans.nl

Abstract

Background: Physical inactivity is a major public health problem. The *It's LiFe!* monitoring and feedback tool embedded in the Self-Management Support Program (SSP) is an attempt to stimulate physical activity in people with chronic obstructive pulmonary disease or type 2 diabetes treated in primary care.

Objective: Our aim was to evaluate whether the SSP combined with the use of the monitoring and feedback tool leads to more physical activity compared to usual care and to evaluate the additional effect of using this tool on top of the SSP.

Methods: This was a three-armed cluster randomised controlled trial. Twenty four family practices were randomly assigned to one of three groups in which participants received the tool + SSP (group 1), the SSP (group 2), or care as usual (group 3). The primary outcome measure was minutes of physical activity per day. The secondary outcomes were general and exercise self-efficacy and quality of life. Outcomes were measured at baseline after the intervention (4-6 months), and 3 months thereafter.

Results: The group that received the entire intervention (tool + SSP) showed more physical activity directly after the intervention than Group 3 (mean difference 11.73, 95% CI 6.21-17.25; P<.001), and Group 2 (mean difference 7.86, 95% CI 2.18-13.54; P=.003). Three months after the intervention, this effect was still present and significant (compared to Group 3: mean difference 10.59, 95% CI 4.94-16.25; P<.001; compared to Group 2: mean difference 9.41, 95% CI 3.70-15.11; P<.001). There was no significant difference in effect between Groups 2 and 3 on both time points. There was no interaction effect for disease type.

Conclusions: The combination of counseling with the tool proved an effective way to stimulate physical activity. Counseling without the tool was not effective. Future research about the cost-effectiveness and application under more tailored conditions and in other target groups is recommended.

Trial Registration: ClinicalTrials.gov: NCT01867970, https://clinicaltrials.gov/ct2/show/NCT01867970 (archived by WebCite at http://www.webcitation.org/6a2qR5BSr).

(J Med Internet Res 2015;17(7):e184) doi:10.2196/jmir.4579

KEYWORDS

motor activity; behavior change; self-management support; primary care nursing; remote sensing technology; COPD; type 2 diabetes

XSL•FO
RenderX

Introduction

Physical inactivity is a major public health problem [1,2] because it increases the risk of several diseases, such as coronary heart disease, type 2 diabetes, and several types of cancer. It also shortens life expectancy [1]. For people with a chronic disease, physical inactivity enhances the chance of complications and comorbidities [3]. Unfortunately, about one-third of adults worldwide do not reach public health guidelines for recommended levels of physical activity (PA) [4]. Therefore, the promotion of PA is a public health priority [5]. One of the approaches to increase PA is through primary health care [6]. Because practice nurses have frequent contact with people with chronic conditions to monitor treatment outcomes, it is recommended that they incorporate support to change physical inactivity behaviors [7,8]. However, providing only verbal advice has proven to be insufficient [9]. Despite the heterogeneity in results of physical activity intervention studies, the most effective approach is professional advice and guidance with continued support and combining a mix of behavior change strategies [10-12]. Effective behavior change strategies for the promotion of PA are self-monitoring, providing feedback for behavior, goal setting, providing tools to facilitate behavior, action planning, social support, barrier identification, and providing information on the consequences specific to the individual [10,11,13].

An example of a tool to facilitate behavior is the use of innovative technology such as mobile phones with built-in, or in combination with, pedometers or accelerometers. These technologies can facilitate self-monitoring, goal setting, and real-time feedback. Despite the fact that general mobile phone use is growing as well as mobile phone use in PA research [14], there is a lack of well-designed experimental studies with appropriate intervention periods and sample sizes [15] to explore whether these technologies add value on top of behavior change counseling by the practice nurse (PN). The *It's LiFe!* intervention is a combination of behavior change strategies delivered by the PN in a self-management support program (SSP) that is partly integrated with usual care as well as the use of a monitoring and feedback tool for patients in daily life.

A cluster randomized controlled trial was conducted to evaluate the longitudinal effects of this multifaceted intervention on 40-70 year-old patients with chronic obstructive pulmonary disease (COPD) and diabetes type 2 (DM2) in primary care. Furthermore, the additional effect of using this tool on top of the SSP was evaluated. The main hypothesis was that after a 4-6 month intervention period, the complete intervention increases participants' moderate to vigorous physical activity by at least 10 minutes per day compared to care as usual and that this increase maintains over 3 months.

Methods

The study methods, intervention, and outcomes have been reported in detail previously [16]. See Multimedia Appendix 1 for the CONSORT-EHEALTH checklist [17].

Study Design

A three-arm clustered randomized controlled trial among 24 general practices in the south of the Netherlands was conducted (NCT01867970). A cluster design was chosen to avoid contamination by unintended influence of the PN in the control group. After stratification based on the number of registered DM2 patients per practice, two blocks of 12 practices were randomly assigned to three groups using sealed envelopes. Practices allocated to Group 1 received the complete intervention (monitoring and feedback tool and SSP), practices in Group 2 received the SSP only, whereas practices in Group 3 received care as usual. Four strata were defined: small (<90 DM2 patients), medium (90-190), large (190-390), and extra-large (>390). There was no blinding for allocation of practices. The research team was blinded for allocation of participants during the analysis phase. Data were analyzed anonymously and coding was revealed after analyses.

Participants: Practices and Patients

We invited 250 family practices in the South of Netherlands by invitation letter, telephone, or personal contact, until 24 practices agreed to participate. Eligibility for participants was determined as follows: between 40 and 70 years old with DM2 or COPD, and who did not, according to the PN, comply with the Dutch Norm for Healthy Exercise (having at least 30 minutes of moderate to vigorous physical activity on 5 or more days of the week) [18]. Additional inclusion criteria for the DM2 patients was a body mass index (BMI) >25, and for the COPD patients, a clinical diagnosis of COPD according to the GOLD-criteria stage 1-3, known to be stable in their respiratory function for at least 6 weeks, and on a stable drug regimen. Furthermore, participants needed to be able to access a computer with an Internet connection and master the Dutch language sufficiently.

Exclusion criteria were the presence of coexisting medical conditions with a low survival rate, severe psychiatric illness, or chronic disorders or diseases that seriously influence the ability to be physically active, and being treated primarily by a medical specialist or participating in another PA intervention. The PNs in each practice were asked to send 20-32 general invitation letters to patients who met the inclusion criteria. After randomization, the PN called the patients to give specific information about the allocated condition and ask if they wanted to participate. If the patient decided to participate, they received a specific information letter and an informed consent form. Each practice was instructed to include 5-7 patients with DM2 and 5-7 patients with COPD. This study was approved by the Medical Ethical Committee of the Maastricht University/Academic Hospital Maastricht in the Netherlands (12-3-071).

Intervention

Overview

The complete *It's LiFe!* intervention consisted of the self-management support program and a monitoring and feedback tool. Both elements were developed in a user-centered design process and tested on usability and feasibility [19-23]. Furthermore, 2 patient representatives from the Netherlands Asthma Foundation and the Dutch Diabetes Association

XSL•FO
RenderX

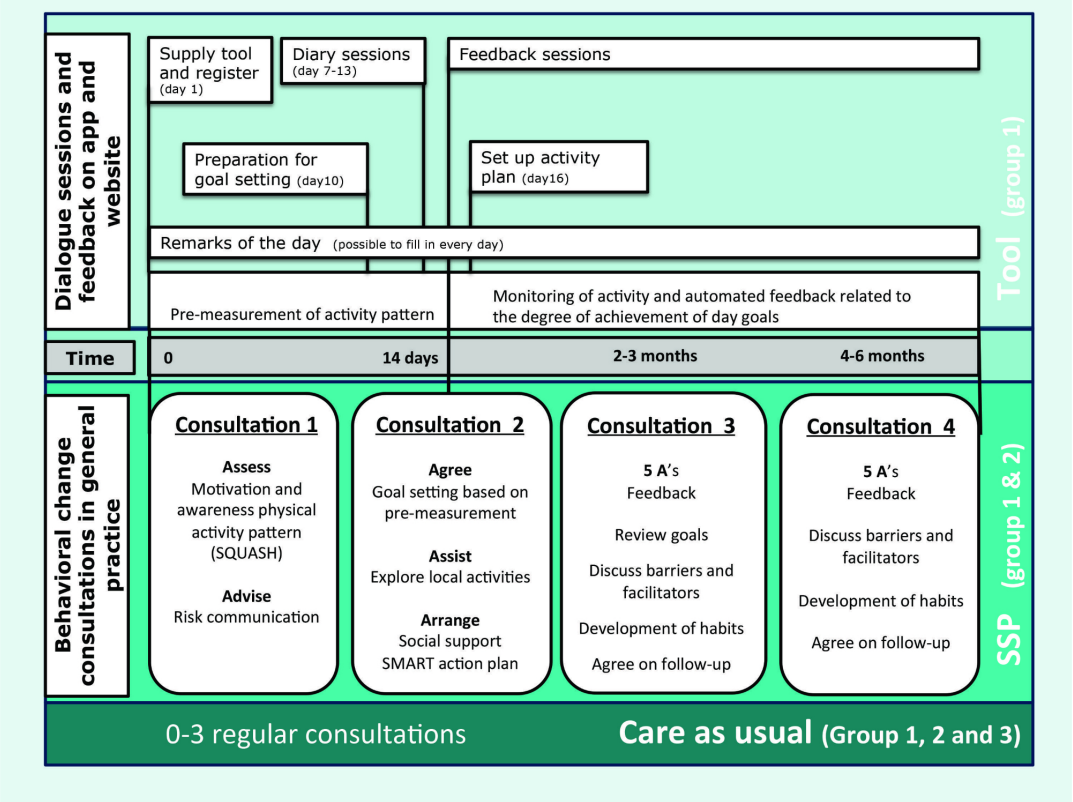
participated in the research group to provide feedback on every aspect of the trial.

The Self-Management Support Program

The program consisted of four individual consultations with the PN; in the first week, after 2 weeks, after 2-3 months, and after 4-6 months (Figure 1) [19]. First, the participants received an information booklet about the course of the intervention containing the Short Questionnaire to Assess Health-Enhancing PA (SQUASH) [24] and a list of locally organized PA activities.

In the first consultation, the PN raised awareness about the risks of physical inactivity, and the PA level of the patient was discussed using the previously completed SQUASH questionnaire. In addition, participants received a general and a disease-specific pamphlet about PA [25-27]. Between the first and the second consultation, a pre-measurement of the activity pattern was taken, and participants answered questions about barriers and facilitators for PA. In Group 1, PA was objectively measured by the tool, and all questions were answered via a dialogue session on the tool. Group 2 kept a PA diary on paper and answered questions about barriers and facilitators in the information booklet. During the second consultation, a personal goal was set in minutes of activity per day based on the pre-measurement, and the PN encouraged the participants to set up an activity plan to reach personal goals. Furthermore, the nurse informed the participants about locally organized PA options. In the third consultation, possibly by mail or telephone, activity results, barriers, facilitators, and the creation of new PA habits were discussed, and some participants reconsidered their activity goal. In the last consultation, activity results, barriers, facilitators, and PA habits were evaluated. Furthermore, how the PN and patient would continue the lifestyle coaching was discussed. The consultations were based on the "Five 'A's Cycle" counseling technique (assess-advise-agree-assist-arrange) [28,29].

Figure 1. Course of the It's LiFe! interventions.



The Tool

The tool consists of a three-dimensional (3D) activity monitor, a mobile phone app, and a Web app (Figure 2) [20]. Participants were asked to wear the activity monitor on a daily basis. They could see their real-time activity results and history in minutes of moderate to vigorous activity on the mobile phone and Web app, in relation to a personal goal. During the pre-measurement, participants participated in dialogue sessions (Figure 1). In the "diary sessions", they were asked about enjoyment and exertion of performed activities. In the "preparation for goal setting", they were asked about barriers and facilitators to exercise. Based on the activity results and the answers in the dialogue sessions, a personal activity goal was set in the second consultation of the SSP. Hereafter, automated feedback messages were sent related to the personal goal. Moreover, the participant was asked in a dialogue session to set up an activity plan to achieve the daily goal. During the entire intervention, activity results and answers to dialogue sessions were visible for the PN on a secured Web app [20,23]. The apps were not changed or updated during the trial (version 2.7). For technical questions and problems with the tool, the participants and PNs could contact a helpdesk during working hours to avoid contact between researchers and participants.

XSL•FO
RenderX

Figure 2. The It's LiFe! activity monitor and mobile phone app.



Training of the Practice Nurses

For mastering the execution of the intervention, practice nurses in Groups 1 and 2 received an online Web lecture and consecutively a personal instruction session at their workplace. In addition, they received on paper, an explanation of the Five A's model, the associated counseling techniques, and detailed instruction charts for each consultation. Nurses in Group 1 were able to try out the tool before the start of the consultations.

Data Collection

All participants received a Personal Activity Monitor AM300 (Pam) [30-32] and questionnaires by regular mail, at baseline (t0), after the intervention at 4-6 months after baseline (t1), and 3 months after the end of the intervention, approximately 9 months after baseline (t2). The last measurement was initially set at 6 months after the intervention, but due to time and money constraints, this could not be realized. The Pam was blinded, which means that participants could not read the display with activity information to prevent any feedback and intervention effect of this measurement.

Outcome Measures

The primary outcome measure was the average minutes per day of PA per patient, measured with the Pam [30-32]. The participants were asked to wear the Pam for 8 consecutive days clipped to their waistband on the hip and to record in a diary the time it was worn. A measurement was considered valid if the tool was worn on ≥5 days for ≥8 hours. Minutes per day were divided in three categories according to metabolic equivalent tasks (METS): light (1.8-2.99 METS), moderate (3-6 METS), and vigorous (>6 METS). The number of minutes of PA in the moderate and vigorous category (≥3 METS) was considered the primary outcome measure because moderate to vigorous activity is recommended by the World Health Organization [33]. Secondary outcome measures were general self-efficacy (general self-efficacy scale) [34], exercise self-efficacy (exercise self-efficacy scale) [35-37], and quality of life (RAND 36) [38,39].

Statistical Analysis

The sample size calculation was based on the primary outcome measure (minutes of moderate to vigorous PA per day). Based on a power of 80%, an alpha of .05 (two-tailed testing), an expected difference between Groups 1 and 3 of 10 minutes of PA per day per participant, and an assumed intraclass correlation between the practices of 0.15, we required 72 participants over eight general practices in each group. A dropout rate of 10% was taken into account, which resulted in a desired number of 80 participants per group.

Intention to treat and per protocol analyses were performed. Participants of the intervention groups were included in the per protocol analysis if they received a minimum of three consultations (75%) spread over at least 3 months based on registration forms of the consultations obtained from the PNs. Participants from all groups were excluded from the per protocol analysis if they did not complete the second measurement (t1). Per protocol analysis were conducted to investigate whether results were different if only participants were included who adhered sufficiently to the interventions.

XSL•FO
RenderX

Normal distribution of the data was checked visually using normal q-q plots and histograms. Outliers were not removed. Continuous variables were presented as means, and standard deviation and categorical variables as numbers and percentages. Differences in baseline characteristics between groups at baseline were investigated with chi-square and analysis of variance (ANOVA). Variables that differed with a P value of .10 or smaller were considered as potential confounders in further analysis. For the RAND 36 outcomes, only the physical component and the mental component were used in further analysis, since the eight subscales strongly correlated. To adjust for the dependency of patients within time and practices (intraclass correlation [ICC]), we used restricted maximum likelihood (REML) multilevel analyses with random intercepts. The differences of the -2 log likelihood and degrees of freedom between models were examined to decide if a one-, two-, or three-hierarchical (time, participants, and general practices) model had to be applied (model selection was performed with a maximum likelihood [ML]). Separate models were set up for each outcome measure, adjusted with Bonferroni correction. The independent variables in each model were two dummy variables indicating the group, with the group of participants receiving care as usual as the reference category, and two dummy variables for time and their interaction effects. In addition, outcome estimates of the multilevel analyses were corrected for baseline and for potential confounders (differences between groups at baseline). Potential confounders were stepwise included in the model if the regression coefficients of time, group, and the interaction of group x time changed by ≥10% on average. To study whether the effects in COPD patients differed from the effects in participants with DM2, a subgroup analysis was done by including interaction effects. Missing values on items in questionnaires were handled according to the questionnaire's analysis manual; missing data in follow-up were not imputed as multilevel analysis accounts for that [40]. All analyses were carried out with IBM Statistical Product and Service Solutions (SPSS) Statistics for Windows, version 22.0.

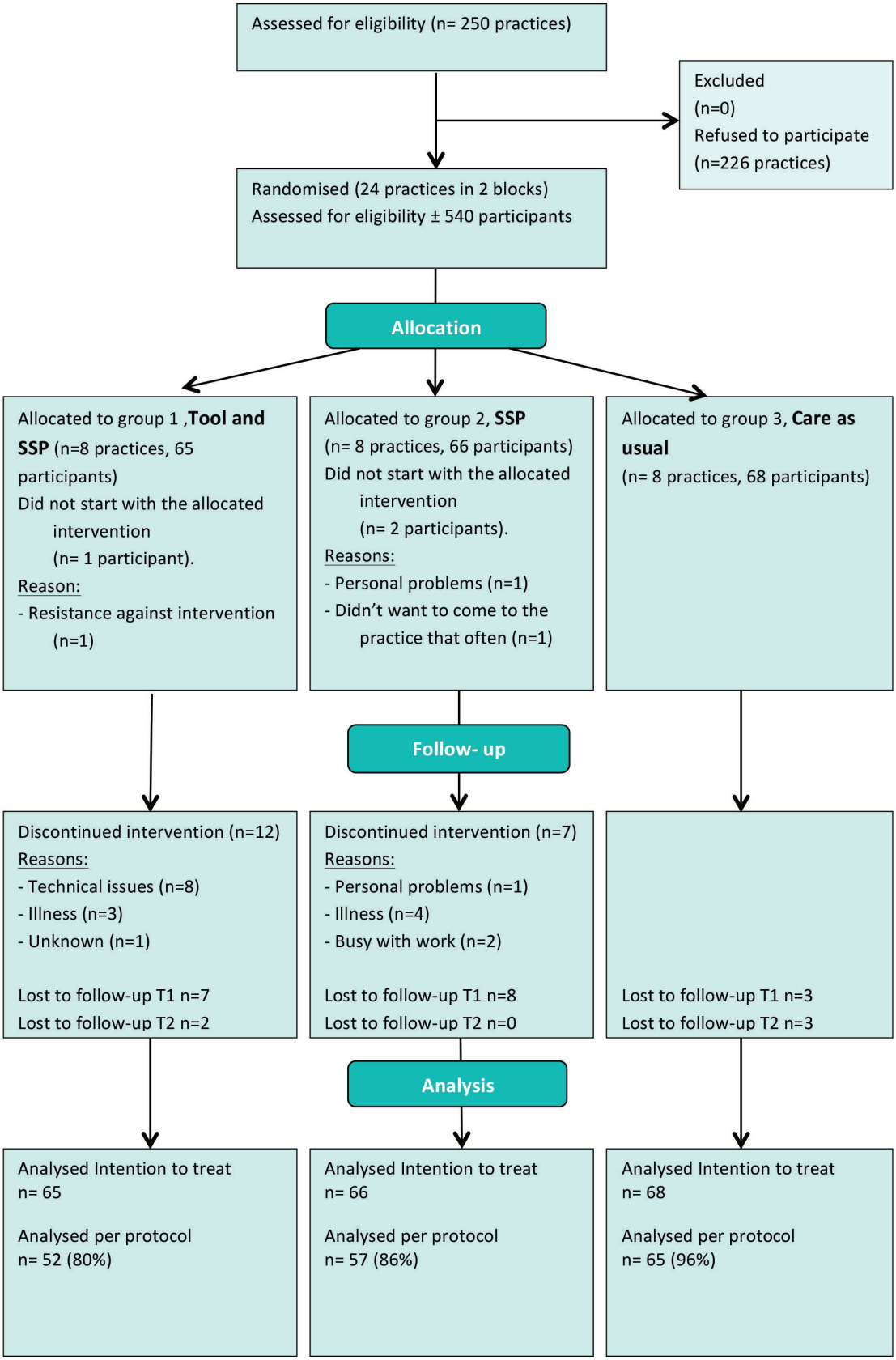
Results

Overview

In total, 24 general practices were randomly assigned to Group 1 (tool and SSP), Group 2 (SSP), or Group 3 (care as usual). In every group, we included one small practice, three medium, three large, and one extra-large practice. The individual practices included 3-14 participants with a median (interquartile range) of 9 participants (7-10 participants). As shown in Figure 3, PNs sent approximately 540 patients a general invitation letter and 199 patients (Group 1: 65 participants, Group 2: 66 participants, Group 3: 68 participants) agreed to participate and completed the baseline measurement. In June 2013, the first practices started with the intervention, and in April 2014 PNs in the last practices performed their last consultations. In Group 1, one participant did not start with the intervention because in his opinion, the intervention was not tailored to his age group, and 12 participants did not receive the minimal intervention as intended. In Group 2, 2 participants dropped out before the start of the intervention and 7 participants did not receive the minimal intervention as intended. In total, 23 participants were lost to follow-up. In the intention-to-treat analyses, data from all participants were taken into account (n=199) (Figure 3). Table 1 shows the baseline characteristics of participants in each group, and Table 2 shows the mean outcome values at baseline. Significant group differences, which were included as confounders in further analyses, were found for BMI, computer use, minutes of PA (≥3 METS), and quality of life (physical component scale).

XSL•FO
RenderX

Figure 3. It's LiFe! CONSORT flow diagram.



XSL•FO
RenderX

Table 1. Baseline characteristics of participants.

Characteristics of participants	Group 1 (n=65), Tool & SSP	Group 2 (n=66), SSP	Group 3 (n=68), Care as usual
Female sex, n (%)	34 (52.3)	31 (47.0)	37 (54.4)
Age in years, mean (SD)	57.5 (7.0)	56.9 (8.3)	59.2 (7.5)
BMI[a], mean (SD)	30.4 (5.7)	29.5 (5.9)	28.2 (4.3)
Origin non-Dutch, n (%)	5 (7.7)	4 (6.1)	3 (4.4)
Married or cohabiting partners, n (%)	48 (73.9)	46 (69.7)	55 (80.9)
Education, n (%)			
Low	19 (29.2)	19 (28.8)	15 (22.1)
Medium	35 (53.8)	40 (60.6)	43 (63.2)
High	11 (16.9)	6 (9.1)	10 (14.7)
Employed, n (%)	31 (47.7)	31 (47.0)	31 (45.6)
COPD, n (%)	25 (38.5)	26 (39.4)	31 (45.6)
Gold stadium, n (%)			
GOLD stadium 1	9 (36.0)	13 (50.0)	15 (48.4)
GOLD stadium 2	15 (60.0)	12 (46.2)	16 (51.6)
GOLD stadium 3	1 (4.0)	1 (3.8)	0 (0.0)
Diabetes type 2, n (%)	40 (61.5)	40 (60.6)	37 (54.4)
Insulin use	3 (7.5)	6 (15.0)	8 (21.6)
Comorbidities, n (%)	51 (78.5)	46 (69.7)	43 (63.2)
Asthma	6 (9.2)	8 (12.1)	4 (5.9)
Cardiac/vascular	12 (18.5)	8 (12.1)	7 (10.3)
Hypertension	22 (33.8)	29 (43.9)	20 (29.4)
Arthritis	13 (20.0)	11 (16.7)	16 (23.5)
Depression	3 (4.6)	5 (7.6)	5 (7.4)
Also diabetes	2 (3.1)	1 (1.5)	1 (1.5)
Also COPD	2 (3.1)	6 (9.1)	2 (2.9)
Other	28 (43.1)	22 (33.3)	27 (39.7)
Computer use [a] , n (%)			
Regularly	50 (76.9)	43 (65.2)	47 (69.1)
Rarely	15 (23.1)	23 (34.8)	21 (30.9)
Mobile phone use, n (%)			
Owns a smartphone	24 (36.9)	24 (36.3)	19 (28.0)
Uses mobile phone frequently	20 (30.8)	20 (30.3)	15 (22.1)
Uses mobile phone rarely	19 (29.2)	19 (28.8)	33 (48.5)
Does not own a mobile phone	2 (3.1)	3 (4.5)	1 (1.5)

[a]$P \leq .10$, tested with chi square or ANOVA.

XSL•FO
RenderX

Table 2. Values at baseline.

	Group 1 (n=65), Tool & SSP	Group 2 (n=66), SSP	Group 3 (n=68), Care as usual
Physical activity, mean (SD)			
Minutes per day in moderate and vigorous ≥3 METS[a]	39.3 (18.1)	47.5 (26.5)	44.1 (20.3)
Wear time of the Pam in hours a day	14.3 (1.7)	14.5 (1.5)	14.3 (1.3)
Self-efficacy, mean (SD)			
General self-efficacy scale	3.2 (0.5)	3.2 (0.5)	3.1 (0.5)
Exercise self-efficacy scale	55.4 (17.0)	53.1 (21.3)	54.0 (19.2)
Factor 1 Situational/interpersonal	51.2 (18.7)	45.9 (20.8)	48.3 (23.2)
Factor 2 Competing demands	62.0 (18.5)	60.0 (21.6)	62.6 (20.2)
Factor 3 Internal feelings	53.8 (18.8)	53.3 (22.2)	52.4 (21.1)
Quality of life, mean (SD)			
Physical Component Score[a]	42.5 (11.1)	46.1 (9.8)	45.8 (9.4)
Mental Component Score	48.2 (10.3)	48.6 (11.7)	50.1 (9.5)
RAND36 physical functioning	68.7 (22.2)	74.6 (20.4)	74.7 (21.9)
RAND36 role functioning physical[b]	55.8 (45.9)	72.2 (36.7)	70.8 (39.5)
RAND36 role functioning emotional	72.8 (38.1)	77.4 (34.4)	78.4 (35.4)
RAND36 social functioning	77.1 (22.8)	77.7 (23.8)	80.5 (20.8)
RAND36 body pain	66.0 (24.8)	70.7 (25.1)	70.8 (23.1)
RAND36 mental health	73.9 (15.1)	74.9 (19.7)	76.5 (14.9)
RAND36 vitality[b]	55.2 (19.1)	62.5 (20.8)	64.3 (16.4)
RAND36 general health	51.3 (19.6)	55.6 (20.6)	55.2 (16.2)

[a]$P \leq .10$, tested with ANOVA.

[b]$P \leq .05$, tested with ANOVA.

Primary Outcome (Intention to Treat)

For the primary outcome, a two-level hierarchical model dealing with dependency of measurements in time within patients (but not family practices) was applied with a correction for baseline physical activity and wear time. ICC for repeated measures was .77, and ICC for participants in the same practice was .005. Directly after the intervention, participants in Group 1 who received the tool and the SSP showed 8 minutes more moderate and vigorous physical activity (≥3 METS) than participants in the SSP, and 12 minutes more PA than the care as usual group. This improvement difference was 9 minutes and 11 minutes, respectively, 3 months after the end of the intervention. No difference was observed between Group 2 (SSP) and Group 3 (care as usual). Results are shown in Table 3.

Table 3. Multilevel analyses for differences between the three groups for physical activity.

Follow-up		Unadjusted mean (SD)			Adjusted mean difference (95% CI); *P* value[a]			ICC[b]
		Tool & SSP	SSP	Care as usual	Tool & SSP: care as usual	SSP: care as usual	Tool & SSP - SSP	
PA moderate and vigorous (≥3METS)[a]	Baseline (t0)	39.29 (18.1)	47.47 (26.5)	44.13 (20.3)	-0.34 (-5.65 to 4.97); 1.000	0.15 (-5.13 to 5.44); 1.000	-0.50 (-5.83 to 4.84); 1.000	.77
	4-6 months (t1)	48.16 (23.8)	46.28 (30.8)	39.61 (19.5)	11.73 (6.21 to 17.25); .000[c]	3.87 (-1.60 to 9.24); .270	7.86 (2.18 to 13.54); .003[c]	
	9 months (t2)	48.82 (23.8)	45.34 (31.3)	42.40 (18.9)	10.59 (4.94 to 16.25); .000[c]	1.19 (-4.38 to 6.76); 1.000	9.41 (3.70 to 15.11); .000[c]	

[a]Adjusted for baseline physical activity and wear time.

[b]2-level random intercept (repeated measurements).

[c]$P < .01$.

XSL•FO
RenderX

Secondary Outcomes

For all secondary outcome measures, a two-level hierarchical model was applied. Table 4 shows that in general and exercise self-efficacy, no significant differences were observed. After 9 months, participants in Group 2 (SSP) did score significantly higher for the physical component of the quality of life scale than participants in Groups 1 (tool + SSP) and 3 (care as usual). At the end of the intervention (6 months), participants in both intervention groups did score significantly higher on the mental component scale compared to the care as usual group.

Table 4. Multilevel analyses for differences between the three groups for secondary outcome measures.

	Unadjusted mean (SD)			Adjusted mean difference (95% CI); *P* value[a]		
Follow-up	Tool +SSP	SSP	Care as usual	Tool +SSP – care as usual	SSP – care as usual	Tool +SSP- SSP
Self-efficacy						
General self-efficacy scale[b]						
Baseline (t0)	3.2 (0.5)	3.2 (0.5)	3.1 (0.5)	0.03 (-0.10 to 0.16); 1.000	0.03 (0.10 to 0.16); 1.000	-0.00 (-0.13 to 0.13); 1.000
4-6 months (t1)	3.3 (0.4)	3.3 (0.5)	3.2 (0.4)	0.05 (-0.09 to 0.18); 1.000	0.02 (-0.11 to 0.15); 1.000	0.03 (-0.10 to 0.16); 1.000
9 months (t2)	3.2 (0.5)	3.3 (0.5)	3.2 (0.4)	0.01 (-0.13 to 0.15); 1.000	0.00 (-0.13 to 0.13); 1.000	0.01 (-0.13 to 0.14); 1.000
Exercise self-efficacy scale[c]						
Baseline (t0)	55.4 (17.0)	53.1 (21.3)	54.0 (19.2)	1.10 (-5.04 to 10.38); 1.000	-0.68 (-8. 36 to 7.01); 1.000	2.67 (-5.04 to 10.38); 1.000
4-6 months (t1)	59.7 (17.3)	59.7 (19.6)	54.5 (17.4)	4.86 (-3.12 to 12.83); .431	5.41 (-2.52 to 13.35); .304	-0.56 (-8.61 to 7.50); 1.000
9 months (t2)	52.1 (16.1)	60.3 (19.1)	56.5 (19.2)	-0.03 (-8.01 to 7.94); 1.000	3.60 (-4.33 to 11.53); .828	-3.63 (-11.69 to 4.43); .838
Quality of life						
RAND physical component[d]						
Baseline (t0)	42.5 (11.1)	46.1 (9.8)	45.8 (9.4)	-0.31 (-2.48 to 1.86); 1.000	0.20 (-1.96 to 2.35); 1.000	-0.51 (-2.69 to 1.68); 1.000
4-6 months (t1)	45.2 (9.5)	46.8 (10.0)	47.0 (10.0)	-0.07 (-2.32 to 2.19); 1.000	-0.08 (-2.33 to 2.17); 1.000	0.01 (-2.30 to 2.33); 1.000
9 months (t2)	44.1 (9.5)	48.2 (8.6)	45.8 (9.5)	0.34 (-1.96 to 2.64); 1.000	2.99 (0.72 to 5.26); 0.005[e]	-2.65 (-4.99 to -0.32); 0.020[f]
RAND Mental component[d]						
Baseline (t0)	48.2 (10.3)	48.6 (11.7)	50.1 (9.5)	-0.30 (-3.27 to 2.68); 1.000	-0.39 (-3.34 to 2.56); 1.000	0.09 (-2.90 to 3.09); 1.000
4-6 months (t1)	48.8 (10.6)	51.6 (11.3)	47.7 (9.8)	3.23 (0.14 to 6.32); 0.04[f]	4.39 (1.32 to 7.47); 0.002[e]	-1.16 (-4.33 to 2.01); 1.000
9 months (t2)	48.3 (11.7)	50.1 (10.9)	50.3 (8.3)	0.21 (-2.94 to 3.36); 1.000	0.23 (-2.88 to 3.34); 1.000	-0.02 (-3.22 to 3.17); 1.000

[a]Linear mixed model 2-level random intercept (repeated measurements).

[b]Adjusted for baseline general self-efficacy scale, computer use, and baseline physical activity moderate + vigorous.

[c]Adjusted for baseline exercise self-efficacy scale.

[d]Adjusted for baseline RAND physical component and baseline RAND mental component.

[e]$P<.01$.

[f]$P<.05$.

Per Protocol Analyses

The results from 174 participants (Figure 3) were analyzed for the per protocol analysis. All per protocol analysis confirmed the intention to treat analysis.

Subgroup Analyses

No differences were observed in outcomes for people with COPD or type 2 diabetes (results not presented).

XSL•FO
RenderX

Discussion

Principal Findings

The complete *It's LiFe!* intervention led to significant improvement of moderate to vigorous physical activity among patients with COPD or type 2 diabetes between 40 and 70 years old in primary care, compared to usual care. Right after the intervention period, the entire intervention added 12 minutes per day of moderate to vigorous physical activity compared to care as usual. Three months after the intervention period, this progress was still significant (11 minutes). This study also proved that use of the tool on top of the SSP is more effective than SSP only. The added value of the tool was an additional 8 minutes of moderate to vigorous physical activity per day. The SSP alone had no significant effect on physical activity compared to care as usual. For the secondary outcome measures, the intervention effect was not evident. It did not result in higher self-efficacy levels. Only the scores on the mental component scale of quality of life showed higher levels directly after both interventions, compared to care as usual, but this difference was not maintained after 9 months. At 9 months follow-up, participants in the SSP group scored significantly higher on the physical component of the quality of life scale compared to the other groups.

Comparison With Prior Work

From the result that the tool embedded in the SSP is effective in contrast to the SSP alone, we can conclude that the automated self-monitoring and feedback component and/or the fact that the PN could see the objective measured PA results, was the most powerful element of the combined intervention. This is in line with the conclusion of a meta-analysis, that PA intervention studies for chronically ill patients incorporating self-monitoring showed a greater effect than studies without self-monitoring [41]. In the SSP, participants only monitored their behavior during the first 2 weeks by using an activity diary. The fact that PA was measured objectively in Group 1 may also have reinforced the goal setting component. Goal setting is more effective if goals are set with a specific outcome, proximal in terms of attainment, and realistic for the individual [13]. This is easier to achieve if objective PA results are available for the patient and the PN, and goals can be adapted during the intervention period based on the obtained results. The individual effect of the tool without the guidance by the PN cannot be extracted from this research, although we do expect that guidance by the nurse is an essential element of the intervention for first raising awareness, risk communication, social support, perseverance with the intervention, and adoption and persistent use of the tool. From the pilot study, we learned that participants felt a desire to succeed due to the commitment they made with the PN and the effort she put into them [22]. Other research also showed the importance of professional advice and guidance with continued support for the improvement of physical activity levels [12].

Other studies demonstrated that a reduction in the number of contacts diminished the behavior change that had been already achieved, especially when the intervention ends [13,42,43]. In this study, 3 months after the intervention period, Group 1 was still significantly more active than the care as usual and the SSP group.

Although exercise self-efficacy is positively correlated with physical activity levels [35], no significant differences were found on this scale between the groups nor on general self-efficacy. This is in line with the findings from the *It's LiFe!* pilot study [22]. Surprisingly, no effects were found on the physical component of the quality of life scale directly after the intervention, but it did improve in the SSP group 3 months after the intervention. We have no explanation for this observation. Awareness that physical activity is being monitored might influence habitual behavior [44]. For the intervention, this was a desirable effect of the *It's LiFe!* tool. However, it was an undesired effect of the use of the Pam. In this view, the proven effectiveness of the total intervention on the primary outcome—moderate to vigorous physical activity—is even more distinct considering the fact that those participating in research often show social desirable behavior while wearing an accelerometer for a short period of time [45]. Participants in Group 1, however, became used to being observed with an accelerometer for 4-6 months, which could have led to less socially desirable behaviors during the research measurement periods, compared to the other groups.

Strengths and Limitations

To the best of our knowledge, this is the first randomized controlled trial that tests the added value of a monitoring and feedback tool in addition to counseling by the PN. An important strength of this study is the objective measurement of the primary outcome measure—physical activity—by an activity monitor instead of a subjective questionnaire. Other strengths are randomization at the practice level to minimize contamination, delay of randomization until after inclusion of the participants, the minimization of Hawthorne effects by avoiding contact between the researchers and participants, and simultaneous with the effect study, a process evaluation was conducted. The latter revealed that despite technical difficulties, the intervention was carried out as intended by the PNs. Another strength of this study is the pragmatic approach. Since the interventions were adapted and embedded in care as usual, it is more likely that the effects will be sustained in the daily primary care setting [46].

Limitations of this study were that the mean baseline physical activity was above the recommended level of 30 minutes of moderate to vigorous activity a day, only 10% of the approached family practices agreed to participate in the study, and only 37% of the approached patients agreed to participate in the intervention. These factors may have induced a selection bias, which makes the results less generalizable. However, a common reason for family practices to refuse participation was the required time investment for the practice nurse. Part of the time investment was for research purposes, which will be eliminated if embedded in daily practice. The low reach among patients may be explained by the fact that in this study patients with low physical activity levels who were not aware of the problem of their inactivity (according to the transtheoretical model of behavior change [47], the precontemplation phase of change) were not included, because the decision to participate had to be

XSL•FO
RenderX

made before the consultation with the PN to create awareness could have taken place. In daily practice, the PN starts with raising awareness in regular consultations, which may result in a shift to the contemplation or preparation phase of change, and after this, patients will be asked if they are willing to work on their lifestyle with the help of the *It's LiFe!* intervention. Another limitation of this study was that cycling, swimming, strength training, and all upper body movements were not taken into account in the primary outcome measure because these could not be captured with the Pam. Furthermore, the follow-up was relatively short—3 months after the intervention period. Ideally, a 12-month follow-up is recommended [48]. Due to time constraints, this was not possible. Clinical outcomes were not measured to avoid the Hawthorne effect in the care as usual group.

Implications for Practice and Future Research

Results of this study revealed the powerful addition of continuous support by the use of a monitoring and feedback tool in addition to behavior change counseling. Because of this added value, it seems worthwhile to implement the intervention on a larger scale. However, cost-effectiveness should be investigated. To encourage general practices to adopt this intervention, health insurance companies should stimulate self-management support regarding physical activity with financial reimbursements for general practices. The fact that the availability and use of smartphones and wearables to measure physical activity is growing [49] is promising for the adoption of the intervention. In daily practice, the intervention can be easily tailored to the individual needs of the patient—for example, more time for raising awareness or referral to an exercise program with a physiotherapist if exercise self-efficacy or capacity is considered too low. In addition, the intervention can be more extensive or recurrent in care as usual with more emphasis on habit formation, instead of a determined period of 4-6 months. The application of this intervention to other target groups should be investigated just as the execution by other care providers as physiotherapists and dieticians.

Conclusions

The monitoring and feedback tool, if embedded into a counseling protocol, was an effective instrument to improve physical activity of patients with COPD or type 2 diabetes between 40 and 70 years old. This improvement was sustained for 3 months. Counseling without the tool was not effective. The use of technology added to counseling is promising for physical activity behavior change. Future research about the cost-effectiveness and application under more tailored conditions and in other target groups is recommended.

Acknowledgments

The project is funded by ZonMw (The Netherlands Organization for Health Research and Development) in the program, "New Instruments for Healthcare", and partly by the insurance company, CZ. The companies involved in the development of the tool were IDEE Maastricht (the Netherlands); Maastricht Instruments Ltd. (the Netherlands); and Sananet Care Ltd. (the Netherlands). During the trial, Sananet Care Ltd. manned the helpdesk and Maastricht Instruments Ltd. sent new tools to the participants if technical problems arose.

The authors would like to thank the participants in the study: the patients, practice nurses and GPs, and the patient representatives, Jos Donkers and Ina van Opstal. Furthermore, we would like to acknowledge April Boessen for her help with the data entry, Claudia Valentijn for her help at the service desk, and Willem Bekker for the 100% physical activity data check.

Conflicts of Interest

A grant from ZonMw and from Insurance Company CZ was received during the conduct of the study.

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.1 [17].

[PDF File (Adobe PDF File), 963KB - jmir_v17i7e184_app1.pdf]

References

1. Lee I, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT, Lancet Physical Activity Series Working Group. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. Lancet 2012 Jul 21;380(9838):219-229 [FREE Full text] [doi: 10.1016/S0140-6736(12)61031-9] [Medline: 22818936]
2. Blair SN. Physical inactivity: the biggest public health problem of the 21st century. Br J Sports Med 2009 Jan;43(1):1-2. [Medline: 19136507]
3. Warburton DER, Nicol CW, Bredin SSD. Health benefits of physical activity: the evidence. CMAJ 2006 Mar 14;174(6):801-809 [FREE Full text] [doi: 10.1503/cmaj.051351] [Medline: 16534088]
4. 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]

XSL•FO
RenderX

5. World Health Organization. Global Strategy on Diet, Physical Activity and Health. Physical Inactivity: A Global Public Health Problem URL: http://www.who.int/dietphysicalactivity/factsheet_inactivity/en/ [accessed 2015-04-15] [WebCite Cache ID 6Xod3cA93]
6. Global Advocacy for Physical Activity (GAPA) the Advocacy Council of the International Society for Physical ActivityHealth (ISPAH). NCD prevention: investments [corrected] that work for physical activity. Br J Sports Med 2012 Aug;46(10):709-712. [doi: 10.1136/bjsm.2012.091485] [Medline: 22869788]
7. Peterson JA. Get moving! Physical activity counseling in primary care. J Am Acad Nurse Pract 2007 Jul;19(7):349-357. [doi: 10.1111/j.1745-7599.2007.00239.x] [Medline: 17680900]
8. Heiligers P, Noordman J, Korevaar J, Dorsman S, Hingstman L, van Dulmen AM, et al. Kennisvraag: praktijkondersteuners in de huisartspraktijk (POH's) klaar voor de toekomst?. URL: http://www.nivel.nl/sites/default/files/bestanden/Rapport-kennisvraag-poh-def.pdf [accessed 2015-04-15] [WebCite Cache ID 6XodTuywa]
9. Lawlor DA, Hanratty B. The effect of physical activity advice given in routine primary care consultations: a systematic review. J Public Health Med 2001 Sep;23(3):219-226 [FREE Full text] [Medline: 11585195]
10. Avery L, Flynn D, van WA, Sniehotta FF, Trenell MI. Changing physical activity behavior in type 2 diabetes: a systematic review and meta-analysis of behavioral interventions. Diabetes Care 2012 Dec;35(12):2681-2689 [FREE Full text] [doi: 10.2337/dc11-2452] [Medline: 23173137]
11. van Achterberg T, Huisman-de Waal GG, Ketelaar NA, Oostendorp RA, Jacobs JE, Wollersheim HC. How to promote healthy behaviours in patients? An overview of evidence for behaviour change techniques. Health Promot Int 2011 Jun;26(2):148-162 [FREE Full text] [doi: 10.1093/heapro/daq050] [Medline: 20739325]
12. Hillsdon M, Foster C, Thorogood M. Interventions for promoting physical activity. Cochrane Database Syst Rev 2005(1):CD003180 [FREE Full text] [doi: 10.1002/14651858.CD003180.pub2] [Medline: 15674903]
13. Artinian NT, Fletcher GF, Mozaffarian D, Kris-Etherton P, Van HL, Lichtenstein AH, American Heart Association Prevention Committee of the Council on Cardiovascular Nursing. Interventions to promote physical activity and dietary lifestyle changes for cardiovascular risk factor reduction in adults: a scientific statement from the American Heart Association. Circulation 2010 Jul 27;122(4):406-441 [FREE Full text] [doi: 10.1161/CIR.0b013e3181e8edf1] [Medline: 20625115]
14. Cowan LT, Van Wagenen SA, Brown BA, Hedin RJ, Seino-Stephan Y, Hall PC, et al. Apps of steel: are exercise apps providing consumers with realistic expectations?: a content analysis of exercise apps for presence of behavior change theory. Health Educ Behav 2013 Apr;40(2):133-139. [doi: 10.1177/1090198112452126] [Medline: 22991048]
15. Bort-Roig J, Gilson ND, Puig-Ribera A, Contreras RS, Trost SG. Measuring and influencing physical activity with smartphone technology: a systematic review. Sports Med 2014 May;44(5):671-686. [doi: 10.1007/s40279-014-0142-5] [Medline: 24497157]
16. Verwey R, van der Weegen S, Spreeuwenberg M, Tange H, van der Weijden T, de WL. A monitoring and feedback tool embedded in a counselling protocol to increase physical activity of patients with COPD or type 2 diabetes in primary care: study protocol of a three-arm cluster randomised controlled trial. BMC Fam Pract 2014;15:93 [FREE Full text] [doi: 10.1186/1471-2296-15-93] [Medline: 24885096]
17. Eysenbach G, Consort- E. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
18. Kemper H, Ooijendijk W, Stiggelbout M. Consensus over de Nederlandse norm voor gezond bewegen. TSG-Tijdschrift voor Gezondheidswetenschappen 2000;78(3):180-183.
19. Verwey R, van der Weegen S, Spreeuwenberg M, Tange H, van der Weijden T, de WL. Upgrading physical activity counselling in primary care in the Netherlands. Health Promot Int 2014 Dec 23:-. [doi: 10.1093/heapro/dau107] [Medline: 25539787]
20. van der Weegen S, Verwey R, Spreeuwenberg M, Tange H, van der Weijden T, de WL. The development of a mobile monitoring and feedback tool to stimulate physical activity of people with a chronic disease in primary care: a user-centered design. JMIR Mhealth Uhealth 2013;1(2):e8 [FREE Full text] [doi: 10.2196/mhealth.2526] [Medline: 25099556]
21. van der Weegen S, Verwey R, Tange HJ, Spreeuwenberg MD, de Witte LP. Usability testing of a monitoring and feedback tool to stimulate physical activity. Patient Prefer Adherence 2014;8:311-322 [FREE Full text] [doi: 10.2147/PPA.S57961] [Medline: 24669188]
22. Verwey R, van der Weegen S, Spreeuwenberg M, Tange H, van der Weijden T, de WL. A pilot study of a tool to stimulate physical activity in patients with COPD or type 2 diabetes in primary care. J Telemed Telecare 2014 Jan;20(1):29-34. [doi: 10.1177/1357633X13519057] [Medline: 24414397]
23. Verwey R, van der Weegen S, Tange H, Spreeuwenberg M, van der Weijden T, de Witte WL. Get moving: the practice nurse is watching you! A case study of the user-centred design process and testing of a web-based coaching system to stimulate the physical activity of chronically ill patients in primary care. Inform Prim Care 2012;20(4):289-298. [Medline: 23890341]
24. Wendel-Vos GC, Schuit AJ, Saris WHM, Kromhout D. Reproducibility and relative validity of the short questionnaire to assess health-enhancing physical activity. J Clin Epidemiol 2003 Dec;56(12):1163-1169. [Medline: 14680666]

XSL•FO
RenderX

25. NISB. 2012. Sportief bewegen met een chronische longaandoening URL: http://www.30minutenbewegen.nl/themas/gezondheid/downloads/sportief-bewegen-met-een-longaandoening.pdf [accessed 2014-04-30] [WebCite Cache ID 6PDsk5Av1]
26. NISB. 2012. Sportief bewegen met diabetes mellitus URL: http://www.sportzorg.nl/_asset/_public/Files/Sportblessures/Diabetes_30minbewegen.pdf [accessed 2014-04-30] [WebCite Cache ID 6PDsgeCvk]
27. NISB. 30 minutenbewegen. 30 Beweegtips URL: http://www.30minutenbewegen.nl/folder-beweegtips.pdf [accessed 2014-04-30] [WebCite Cache ID 6PDsoFdJt]
28. Glasgow RE, Emont S, Miller DC. Assessing delivery of the five 'As' for patient-centered counseling. Health Promot Int 2006 Sep;21(3):245-255 [FREE Full text] [doi: 10.1093/heapro/dal017] [Medline: 16751630]
29. Whitlock EP, Orleans CT, Pender N, Allan J. Evaluating primary care behavioral counseling interventions: an evidence-based approach. Am J Prev Med 2002 May;22(4):267-284. [Medline: 11988383]
30. Damen E. Secondary Pam. 2014. Pam URL: http://www.pam.com/index.php?pid=1 [accessed 2014-05-01] [WebCite Cache ID 6PEz8ql7Y]
31. Slootmaker SM, Chin AP, Schuit AJ, van MW, Koppes LLJ. Concurrent validity of the PAM accelerometer relative to the MTI Actigraph using oxygen consumption as a reference. Scand J Med Sci Sports 2009 Feb;19(1):36-43. [doi: 10.1111/j.1600-0838.2007.00740.x] [Medline: 18266793]
32. Vooijs M, Alpay LL, Snoeck-Stroband JB, Beerthuizen T, Siemonsma PC, Abbink JJ, et al. Validity and usability of low-cost accelerometers for internet-based self-monitoring of physical activity in patients with chronic obstructive pulmonary disease. Interact J Med Res 2014;3(4):e14 [FREE Full text] [doi: 10.2196/ijmr.3056] [Medline: 25347989]
33. World Health Organization. 2011. Global Recommendations on Physical Activity for Health URL: http://www.who.int/dietphysicalactivity/physical-activity-recommendations-18-64years.pdf?ua=1 [accessed 2014-05-01] [WebCite Cache ID 6PF1UBuEa]
34. Schwarzer R, Jerusalem M. Generalized self-efficacy scale. In: Weinman J, Wright S, Johnston M, editors. Measures in health psychology: a user's portfolio. Causal and control beliefs. Windsor, UK: NFER-NELSON; 1995:35-37.
35. van der Heijden MMP, Pouwer F, Pop VJM. Psychometric properties of the Exercise Self-efficacy Scale in Dutch Primary care patients with type 2 diabetes mellitus. Int J Behav Med 2014 Apr;21(2):394-401. [doi: 10.1007/s12529-013-9308-z] [Medline: 23550033]
36. Shin Y, Jang H, Pender NJ. Psychometric evaluation of the exercise self-efficacy scale among Korean adults with chronic diseases. Res Nurs Health 2001 Feb;24(1):68-76. [Medline: 11260587]
37. Everett B, Salamonson Y, Davidson PM. Bandura's exercise self-efficacy scale: validation in an Australian cardiac rehabilitation setting. Int J Nurs Stud 2009 Jun;46(6):824-829. [doi: 10.1016/j.ijnurstu.2009.01.016] [Medline: 19261281]
38. Van der Zee K, Sanderman R. RAND-36. Vol 2. Groningen, The Netherlands: Groningen Research Institue Share; Jun 2012:2012.
39. Ware J, Kosinski M, Keller S. SF-36 physical and mental health summary scales: a user's manual:. In: Health Assessment Lab. Boston: The Health Institute; 1994.
40. Twisk J. Applied multilevel analysis: a practical guide. Cambridge, UK: Cambridge University Press; 2006.
41. Conn VS, Hafdahl AR, Brown SA, Brown LM. Meta-analysis of patient education interventions to increase physical activity among chronically ill adults. Patient Educ Couns 2008 Feb;70(2):157-172 [FREE Full text] [doi: 10.1016/j.pec.2007.10.004] [Medline: 18023128]
42. Yeh M, Rodriguez E, Nawaz H, Gonzalez M, Nakamoto D, Katz DL. Technical skills for weight loss: 2-y follow-up results of a randomized trial. Int J Obes Relat Metab Disord 2003 Dec;27(12):1500-1506. [doi: 10.1038/sj.ijo.0802430] [Medline: 14634681]
43. Thomas SA, Browning C. Behavioural change: an evidence-based handbook for social and public health. Edinburgh: Elsevier/Churchill Livingstone; 2005.
44. Pedišić Ž, Bauman A. Accelerometer-based measures in physical activity surveillance: current practices and issues. Br J Sports Med 2015 Feb;49(4):219-223. [doi: 10.1136/bjsports-2013-093407] [Medline: 25370153]
45. Casaburi R. Activity monitoring in assessing activities of daily living. COPD 2007 Sep;4(3):251-255. [doi: 10.1080/15412550701480158] [Medline: 17729069]
46. Linmans JJ, Viechtbauer W, Koppenaal T, Spigt M, Knottnerus JA. Using electronic medical records analysis to investigate the effectiveness of lifestyle programs in real-world primary care is challenging: a case study in diabetes mellitus. J Clin Epidemiol 2012 Jul;65(7):785-792. [doi: 10.1016/j.jclinepi.2012.01.010] [Medline: 22560776]
47. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot 1997;12(1):38-48. [Medline: 10170434]
48. Foster C, Richards J, Thorogood M, Hillsdon M. Remote and web 2.0 interventions for promoting physical activity. Cochrane Database Syst Rev 2013 Sep 30;9:CD010395. [doi: 10.1002/14651858.CD010395.pub2] [Medline: 24085594]
49. Krijgsman J, Peeters J, Burghouts A, Brabers A, de JJ, Beenkens F. Op naar meerwaarde! eHealth-monitor. Den Haag and Utrecht, The Netherlands: Nictiz and het NIVEL; Oct 2014:9789082030419.

XSL•FO
RenderX

Abbreviations

BMI: body mass index
COPD: chronic obstructive pulmonary disease
DM2: type 2 diabetes
GP: general practitioner
***It's LiFe!*:** Interactive Tool for Self-management through Lifestyle Feedback
METS: metabolic equivalents
ML: maximum likelihood
PA: physical activity
Pam: Physical Activity Monitor AM300
PN: practice nurse
REML: restricted maximum likelihood
SSP: self-management support program

Edited by G Eysenbach; submitted 26.04.15; peer-reviewed by J West, L Nemeth; comments to author 23.06.15; revised version received 04.07.15; accepted 07.07.15; published 24.07.15.

Please cite as:
van der Weegen S, Verwey R, Spreeuwenberg M, Tange H, van der Weijden T, de Witte L
It's LiFe! Mobile and Web-Based Monitoring and Feedback Tool Embedded in Primary Care Increases Physical Activity: A Cluster Randomized Controlled Trial
J Med Internet Res 2015;17(7):e184
URL: http://www.jmir.org/2015/7/e184/
doi:10.2196/jmir.4579
PMID:26209025

©Sanne van der Weegen, Renée Verwey, Marieke Spreeuwenberg, Huibert Tange, Trudy van der Weijden, Luc de Witte. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 24.07.2015. 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 Web-Based Adolescent Positive Psychology Program in Schools: Randomized Controlled Trial

Rowan Burckhardt[1], MPsych(Clinical); Vijaya Manicavasagar[1], PhD; Philip J Batterham[2], PhD; Leonie M Miller[3], PhD; Elizabeth Talbot[1], BPsych(Hons); Alistair Lum[1], BPsych(Hons)

[1]University of NSW and the Black Dog Institute, Randwick, Australia

[2]National Institute for Mental Health Research, The Australian National University, Canberra, Australia

[3]School of Psychology, University of Wollongong, Wollongong, Australia

Corresponding Author:
Rowan Burckhardt, MPsych(Clinical)
University of NSW and the Black Dog Institute
Black Dog Institute
Hospital Rd
Randwick, 2031
Australia
Phone: 61 293824530
Fax: 61 293828501
Email: r.burckhardt@unsw.edu.au

Abstract

Background: Adolescent mental health is characterized by relatively high rates of psychiatric disorders and low levels of help-seeking behaviors. Existing mental health programs aimed at addressing these issues in adolescents have repeated inconsistent results. Such programs have generally been based on techniques derived from cognitive behavioral therapy, which may not be ideally suited to early intervention among adolescent samples. Positive psychology, which seeks to improve well-being rather than alleviate psychological symptoms, offers an alternative approach. A previous community study of adolescents found that informal engagement in an online positive psychology program for up to 6 weeks yielded significant improvements in both well-being and depression symptoms. However, this approach had not been trialed among adolescents in a structured format and within a school setting.

Objective: This study examines the feasibility of an online school-based positive psychology program delivered in a structured format over a 6-week period utilizing a workbook to guide students through website content and interactive exercises.

Methods: Students from four high schools were randomly allocated by classroom to either the positive psychology condition, "Bite Back", or the control condition. The Bite Back condition consisted of positive psychology exercises and information, while the control condition used a series of non-psychology entertainment websites. Both interventions were delivered online for 6 hours over a period of 4-6 weeks during class time. Symptom measures and measures of well-being/flourishing and life satisfaction were administered at baseline and post intervention.

Results: Data were analyzed using multilevel linear modeling. Both conditions demonstrated reductions in depression, stress, and total symptom scores without any significant differences between the two conditions. Both the Bite Back and control conditions also demonstrated significant improvements in life satisfaction scores post intervention. However, only the control condition demonstrated significant increases in flourishing scores post intervention.

Conclusions: Results suggest that a structured online positive psychology program administered within the school curriculum was not effective when compared to the control condition. The limitations of online program delivery in school settings including logistic considerations are also relevant to the contradictory findings of this study.

Trial Registration: Australian New Zealand Clinical Trials Registry: ACTRN1261200057831; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=362489 (Archived by Webcite at http://www.webcitation.org/6NXmjwfAy).

(J Med Internet Res 2015;17(7):e187) doi:10.2196/jmir.4329

XSL•FO
RenderX

KEYWORDS

adolescent; Internet; early medical intervention; randomized controlled trial

Introduction

Epidemiological data indicate that 1 in 4 young people in Australia between the ages of 16-24 years have experienced at least one mental disorder in the preceding year and that young people have the highest rates of mental disorders compared to any other age group [1]. It has been estimated that of those aged 16-24 years experiencing mental health issues, 77% did not access any medical or professional services in the preceding 12 months and that they account for almost 30% of the total mental health burden in Australia [2,3]. This is further aggravated by their relative lack of skills in dealing with life stresses and emotional distress compared to adults [4].

A number of early intervention programs have been developed to address skills deficits and to build resilience in young people. Most of these programs have used cognitive behavioral therapy techniques with varying degress of success [5]. This suggests that researchers need to look further afield for effective prevention approaches to youth mental health such as the building of skills and strategies to enhance personal strengths [6].

In this regard, positive psychology—the scientific study of happiness, well-being, and flourishing—may offer an alternative to conventional prevention programs for youth. Positive psychology can be conceptualized as a range of behavioral, cognitive, and emotional domains that are important for improving well-being. These include developing healthy social relationships, becoming more optimistic, finding "meaning", and practicing mindfulness. Evaluations of non-school-based programs have reported on the benefits of engaging in positive psychology exercises. For example, interventions that increase hope have been shown to predict lower illicit substance use; lower levels of depression, anxiety, and hostility; fewer behavioral problems; and higher academic performance in adolescents [7,8].

Manicavasagar et al conducted a feasibility study of a community-based online positive psychological program, "Bite Back", for adolescents aged 12-18 years [9,10]. Bite Back consists of information and interactive activities relating to nine domains: gratitude, optimism, flow, meaning, hope, mindfulness, character strengths, healthy lifestyle, and positive relationships. Participants in both the Bite Back condition and a control condition were encouraged to regularly use their respective websites over a 6-week period. Symptom measures of depression, anxiety and stress, and well-being were compared within and between conditions and at baseline and post intervention. At the start of the intervention, participants reported low levels of psychopathology, as was expected in a non-clinical sample. Results demonstrated significant decreases in symptom scores and increased scores on well-being post intervention for the Bite Back relative to the control condition. It was reported that when participants with high levels of engagement were examined in isolation, the findings were even stronger suggesting that increased usage was related to greater benefits. Furthermore, qualitative analyses indicated that Bite Back was well received by participants allocated to this condition. It is notable that participants using this online positive psychology program could freely engage with the material at their own pace and in their own time. However, the open-access online format of this program meant that the researchers were unable to examine how participants navigated the website and whether they were exposed to the full range of material and information about positive psychology.

A literature search did not uncover any previously evaluated computer-based school-delivered positive psychology interventions. However, group-format school-based positive psychology interventions, such as the Positive Psychology Program and Wellness-Promotion Intervention, consist of lesson-based modules pertaining to specific topics and skills delivered over the intervention period [11-16]. It was therefore deemed timely to evaluate Bite Back in a structured lesson-based format that was appropriate for delivery within a school environment. This study explores the feasibility of implementing a structured workbook-guided version of Bite Back in senior schools as part of the schools' curriculum. It was expected that the benefits found in our previous study would be extended to this student population, specifically, that participants would improve on measures of well-being and report decreased symptoms scores following the intervention compared to a control condition.

Methods

Participants

Four Australian high schools agreed to participate in the study: two Anglican girls' schools (referred to as "Girls School A" and "Girls School B"), a Catholic boys' school (referred to as "Boys School"), and a Jewish co-educational school (referred to as "Co-ed School"). The level of religiosity of students within these schools was not formally assessed although they are considered by reputation to be approximately equally based on conversations with staff in the educational system. Students were in Grades 7 through 12, which in the Australian educational system are the final 6 years before being eligible to go to university. These four schools were among the highest in terms of socioeconomic status compared to other schools in Australia. Students were included in the study only if they completed the relevant consent forms. This study was approved by the University of NSW Human Research Ethics Advisory Committee (Ethics Approval number 2011-7-35). See Multimedia Appendix 1 for the CONSORT-EHEALTH checklist [17].

Interventions

The Positive Psychology Condition (Bite Back)

Bite Back was developed by the Black Dog Institute to improve the well-being and happiness of young Australians aged 12-18 years. Key objectives of this program are to encourage young people to work to their full potential, become more fully engaged

XSL•FO
RenderX

in all aspects of their lives, and ultimately, to build resilience. The program consists of a range of interactive activities including making gratitude entries, mindfulness meditations, describing personal stories, and a mindfulness exercise involving taking photos. Figure 1 provides screenshots of the website to illustrate key components. Online interactive exercises are designed to encourage generalization of online activities to the "real world", so that young people benefit from implementing the central tenets of positive psychology into their daily lives. Also included on the site is information about nine positive psychology domains and ways to engage with them outside of the website: Gratitude, Optimism, Flow, Meaning, Hope, Mindfulness, Character Strengths, Healthy Lifestyle, and Positive Relationships. The nine domains were based on research that suggested they were important in improving well-being [18]. Users are able to comment on activities (such as submitted photos or stories by other users) with their anonymous profile thereby enabling participation without fear of stigma or judgment. The website is pre-moderated by website staff to ensure that antisocial or concerning behavior is addressed prior to possible posting on the website. In the rare cases that a posting is either concerning from a mental health perspective or abusive, a senior clinical psychologist advised on how to handle it (eg, send an email suggesting they seek external assistance). Confidentiality of users was ensured by requiring new users to sign up using a pseudonym—no real names are permitted. Users can also choose to post content in private, where they are the only ones who can view it, or they can post it publicly. To ensure security of content and users, Bite Back uses password-protected accounts, encrypted password storage, encrypted login details, and secure external servers. For the purposes of the study, a workbook was developed to guide students through the positive psychology exercises on the website and to pose reflective questions about positive psychology. This ensured that participants had sufficient material to use during the course of the program and encouraged users to engage with several domains within the website. Participants were required to record all comments and answers to questions in their workbooks.

The Control Condition

Participants assigned to the control condition used a workbook that was structured in a similar format to that of the Bite Back workbook but with questions and links to non-psychology websites: Australian Broadcasting Corporation, Australian Youth Climate Coalition, the Natural Resource Defense Council, World Wildlife Fund, Wikipedia, Internet Movie Database, and World Youth News. Questions such as "Do you read online news websites in your own time? If yes, which ones?", "Do you think the news makes a difference to how you live in your daily life? Why?", and "Write a summary of what the article was about" were designed to encourage users to engage with these websites and record their comments and answers to questions in their workbooks.

XSL•FO
RenderX

Figure 1. Screenshots from the Bite Back website.



Procedure

Between January and March 2012, 20 schools were invited to participate in this study, of which four agreed. The 20 schools were chosen based on two criteria: physical proximity to the research institute and being non-governmental. The proximity factor was to facilitate the practical running of the study given that trips to the school were necessary and the non-governmental requirement because access to these schools is less onerous and more timely than governmental institutions, which require

XSL•FO
RenderX

permission from the state school department that takes over 6 months to obtain. Prior to start of the study, the lead author met with senior school staff and welfare officers to convey the requirements of the study and to instruct teachers in the study methodology. Parental and student self-consent were obtained for all participants under the age of 16 years, and student self-consent only was obtained for participants aged 16 years and over. The grade level designated to participate in the study was chosen by the school, and no exclusion criteria were applied.

Teachers were instructed to explain to students that they were participating in a study designed to examine the way that websites affect how young people view the world. Teachers were also required to provide a 5-minute discussion on this topic prior to the beginning of the study and were asked not to mention that the study included a positive psychology intervention. Teachers were responsible for handing out workbooks, providing students with access to relevant websites, and managing students' behavior during class time. Teachers were unaware of which classes had been allocated to the intervention or control conditions. Students had no face-to-face contact with any of the researchers prior to, during, or following the study.

Students were randomized by blocks (classes) using an Excel random number generator to allocate them to either the Bite Back or control conditions by an independent researcher not associated with the data collection. Equivalent numbers of students were allocated to each condition across each grade level and for each school (eg, Grade 9 at the Boys School). In cases where there were uneven numbers of blocks, students were assigned to the Bite Back condition. Students completed an online battery of self-report questionnaires twice during class time: pre-intervention and post intervention.

All schools were instructed to use the workbook over a 6-week period. However, two schools (the Boys School and the Co-ed School) completed the workbook over a 4-week period due to class-time constraints. Teachers were told that a total of 6 hours of face-to-face time was required from students over a period of 4-6 weeks. After each session, students in both the control and Bite Back conditions were asked to email completed sections of their workbooks to the researchers. This strategy was employed to encourage engagement and increase compliance. It also served as an indicator of student adherence to the research tasks. Students were advised that their teachers did not have access to their workbooks and completed questionnaires.

Measures

The Depression, Anxiety, and Stress Scale—Short form (DASS-21) consists of three symptom-based subscales [19]. Each subscale has 7 items that participants respond to on a 4-point Likert scale (0="not at all" to 3="most of the time"). Summed scores for each scale range from 0-42 following conversion of scores to match the DASS-42; more severe symptoms are indicated by higher scores. The DASS has also been demonstrated to correlate closely with the Diagnostic and Statistical Manual of Mental Disorders' anxiety and depressive diagnoses [20]. Although the DASS-21 has been previously used on adolescent samples [21-23], we modified some of the wording to improve comprehension (eg, "I felt down-hearted and blue" was changed to "I felt down"). These changes were approved by the DASS-21 authors and the original meaning of the items remained unchanged. Cronbach alpha was calculated for the sample, which indicated that it was .91 for the depression scale, .79 for anxiety, .81 for stress, and .92 for the total score. The DASS-21 total score was analyzed in addition to subscale scores because of evidence that adolescent psychometrics may not support a three-factor model but rather a single factor [21,23].

Student Life Satisfaction Scale (SLSS) comprises 5 items from the original 7-item SLSS, which are rated on a 7-point Likert scale (1="strongly disagree" to 7="strongly agree") that measures quality of life (eg, "I have what I want in life", "My life is going well") [24]. Higher scores indicate a higher subjective satisfaction with life. Cronbach alpha in the current sample was .90.

The Short Warwick-Edinburgh Mental Well-Being Scale (SWEMWBS) is a 7-item measure that assesses positive mental health (well-being) over the past 2 weeks rated on a 5-point scale (1="none of the time" to 5="all of the time") [25]. Psychometric data of the measure on the original WEMWBS for adolescents indicated satisfactory to high internal consistency (r=.87), and the short version has acceptable test-retest reliability (r=.66, 95% CI 0.59-0.72) [25,26]. In our sample, Cronbach alpha was .82.

Statistical Analysis

Data were analyzed using SPSS version 22 [27]. Independent samples t tests were used to examine group differences at baseline. Multilevel linear modeling was used to examine the effectiveness of the intervention in reducing depression, anxiety and stress, and increasing life satisfaction and well-being. This type of model represents an intention-to-treat analysis under the missing-at-random assumption. Specifically, the effect of time (posttest vs pretest) was estimated using a repeated-measures mixed-effects model, which also accounted for clustering of students within schools using a random intercept for each school. An unstructured variance-covariance structure was assumed for both the repeated effect of time and the random effect of school. Degrees of freedom were estimated using Satterthwaite's correction [28].

Results

Flow of Participants and Participant Characteristics

Figure 2 demonstrates the flow of participants during the course of the study. Four schools provided a total of 572 students across grades 7-12, who were randomly allocated through classroom blocks to either the control or Bite Back conditions. Of those allocated, 338 students completed a battery of online questionnaires at baseline and were considered to be enrolled in the study. From the 234 students who did not complete the baseline questionnaire, 110 encountered technical difficulties that prevented them from being able to access the questionnaire successfully. The other 124 either chose not to fill out the questionnaire or were absent. Of the four schools that were

XSL•FO
RenderX

enrolled, one withdrew from the study (Girls School B) and one ceased to participate because the teacher involved in its implementation went on extended leave (Co-ed School). Baseline characteristics of participants are presented in Table 1.

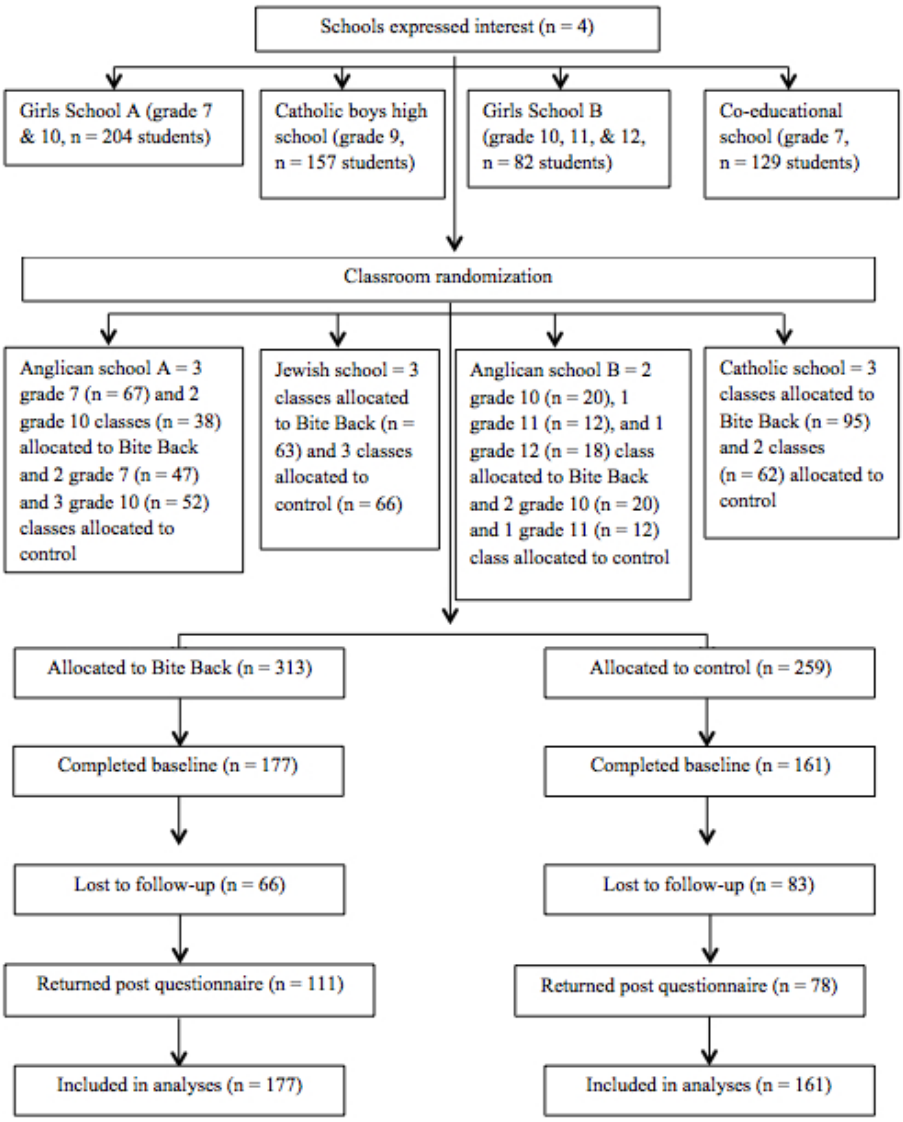
A series of independent samples *t* tests examined baseline differences on questionnaire scores between the two conditions (control versus Bite Back groups across the entire sample). No significant differences were observed between the groups for the depression subscale, P=.40; the anxiety subscale, P=.60; the stress subscale, P=.11; the DASS-total score, P=.25; or the SLSS, P=.15. A significant difference was observed on the SWEMWBS (t_{332}=-4.1, P<.001) indicating that the Bite Back group began the study with higher levels of well-being (mean 18.2) than the control group (mean 16.7). The size of the difference was moderate using the guidelines proposed by Cohen [29] (eta squared=0.05).

Table 1. Baseline characteristics of participants in the school study.

	Girls A		Boys		Girls B		Co-ed	
	Bite Back (n=94)	Control (n=90)	Bite Back (n=72)	Control (n=54)	Bite Back (n=8)	Control (n=5)	Bite Back (n=1)	Control (n=12)
Grades	7 and 10		9		10 and 11		7	
Female, n (%)	184 (100)		0 (0)		13 (100)		3 (23)	
Age in years, mean (SD)	13.8 (1.6)		14.6 (0.4)		15.6 (0.8)		12.6 (0.3)	
Baseline scores, mean (SD)								
DASS depression	9.2 (10.4)	10.2 (9.6)	6.8 (7.1)	8.0 (8.0)	6.0 (3.0)	6.6 (3.2)	—	5.5 (6.6)
DASS anxiety	7.7 (6.9)	7.6 (5.8)	5.0 (5.6)	6.3 (5.8)	4.3 (4.2)	4.4 (2.6)	—	4.4 (2.2)
DASS stress	13.1 (8.6)	13.4 (6.1)	7.8 (5.9)	10.7 (8.0)	10.6 (4.0)	10.4 (4.3)	—	9.9 (6.2)
DASS total	30.2 (22.4)	31.1 (18.0)	19.7 (15.9)	25.0 (18.5)	20.9 (8.2)	21.3 (9.0)	—	19.4 (12.5)
Satisfaction (SLSS)	22.4 (6.4)	21.6 (6.2)	23.6 (4.8)	22.3 (6.0)	25.5 (2.3)	21.6 (6.7)	—	24.9 (4.6)
Flourishing (SWEMWBS)	17.4 (3.6)	16.0 (3.4)	19.1 (2.9)	17.7 (3.2)	19.0 (1.3)	17.6 (1.5)	—	17.0 (3.7)

XSL•FO
RenderX

Figure 2. CONSORT diagram of participants.



Comparing Dropouts to Completers

Participants who dropped out were compared to completers on a series of independent-groups *t* tests. No significant differences were observed on the scales of depression (*P*=.19), anxiety (*P*=.66), stress (*P*=.85), DASS-total (*P*=.59), or the SWEMWBS (*P*=.05). A small difference (eta squared=0.01) was observed on the SLSS (t_{321}=-2.0, *P*=.04), with dropouts reporting significantly less satisfaction (mean 21.8) than completers (mean 23.1).

The two schools that dropped out of the study provided reasons for ceasing to participate. The Anglican school B reported that the "overwhelming feedback from students had been negative" with the students finding both the control and Bite Back conditions "boring". Students reported to their teachers that they did not seem to understand why they were doing this work, and more specifically, what benefits they would derive from their participation. They also reported that they found using the workbooks tedious and laborious. The Co-ed School ceased to participate in the study because the person in charge of implementing the study left 2 weeks after the study commenced and the other teachers did not continue with the research protocol.

Adherence

Adherence was defined as the frequency with which students returned their completed workbook sections (ranging from 0-6): of 336 participants, 27 (8.0%) did not return any of their workbook sections, 52 participants (15.5%) returned 1-2 workbooks, 70 participants (20.8%) returned 3-4 workbooks, and 187 (55.6%) returned 5-6 workbooks.

Due to differences observed in the previous community Bite Back study on levels of engagement with the program, the current data were examined to determine if level of engagement affected outcomes. The sample was dichotomized according to the number of workbook sections completed. Participants who completed either 5 or 6 sections were considered "highly engaged", while those who returned up to 4 completed sections of the workbook were categorized as "low engaged". To ensure that this analysis was not affected by differential baseline scores, *t* tests were used to examine baseline differences on questionnaire scores between participants in the high versus

XSL•FO
RenderX

low engagement groups. No significant differences between the groups were found.

Universal Effects

Overview

To examine the effects of Bite Back on student well-being, an intent-to-treat analysis was completed using multilevel model analysis. Analyses of the universal effects were repeated with only the participants in the high engagement group.

Depression

Results indicated a non-significant interaction between time and condition for DASS-Depression scores: $F_{1,199.9}$=.42, P=.52. The effect of time alone was significant, $F_{1,309.1}$=8.10, P=.005, indicating a significant reduction of depression from baseline (mean 8.57) to post intervention (mean 6.93) for both conditions but no significant differences between the groups. Given that the baseline levels of psychopathology were low, the impact of the website on those reporting high levels of symptoms at the start of the program was also examined. The analysis was repeated among participants in the mild to extremely high categories of DASS depression baseline scores (n=226). No significant differences between the control and the Bite Back conditions were found. To examine if the level of engagement modified the results, the analysis was repeated to compare participants with low versus high engagement with the workbook sections. No significant differences were found between the groups.

Anxiety

There was no significant effect for DASS-Anxiety on the interaction of condition and time. The main effect for time was not significant. No significant interaction between condition and time was found among those with high baseline psychopathology scores, nor the highly engaged group.

Stress

The interaction between time and condition for DASS-Stress scores was not significant. However, the main effect for time was significant, $F_{1,188.4}$=6.19, P=.01, suggesting a reduction in stress scores from baseline (mean 11.47) to post intervention (mean 10.37) for the entire sample. No significant time by condition interactions were observed among participants with high baseline scores. When the highly engaged participants were analyzed, a significant interaction of time and condition was observed: $F_{1,127.9}$=5.56, P=.02. The highly engaged participants in the control group reported a reduction in stress from baseline (mean 12.44) to post intervention (mean 10.01), while the Bite Back participants reported a small increase from baseline (mean 9.49) to post intervention (mean 10.05).

DASS-Total

When total DASS scores were examined, the interaction between time and condition was not significant. However, there was a significant effect of time across both groups, $F_{1,184.57}$=6.67, P=.01, with reductions in total DASS scores from baseline (mean 26.69) to post intervention (mean 23.65) for both conditions. When participants with high baseline scores were analyzed alone, this interaction was not significant: $F_{1,57.75}$=.29, P=.60. For participants who were highly engaged, the interaction of condition and time was not significant: $F_{1,126.47}$=3.11, P=.08.

Life Satisfaction

The interaction of time and condition for the SLSS was not significant. However, the main effect of time was significant, $F_{1,247.0}$=17.35, P<.001, with participants reporting a reduction of life satisfaction from baseline (mean 22.56) to post intervention (mean 21.58). When highly engaged participants were analyzed, the time by condition interaction was significant: $F_{1,153.22}$=6.57, P=.01. Highly engaged control participants reported little change in life satisfaction (mean 21.9 at baseline compared to mean 22.09 at post intervention), while the Bite Back participants reported a larger reduction from baseline (mean 23.99) to post intervention (mean 21.90).

Flourishing

The interaction between time and condition on the SWEMWBS was significant: $F_{1,202.9}$=5.88, P=.02. Baseline scores of participants in the control condition (mean 16.68) increased at post intervention (mean 18.19), while the Bite Back group reported little change (mean 18.18 at baseline increasing to mean 18.37 at post intervention). A significant difference in the interaction of time and condition was also observed when the group with high engagement was analyzed separately: $F_{1,150.4}$=4.15, P=.04.

Discussion

Principal Findings

The aim of the current study was to evaluate the feasibility and effectiveness of the Bite Back program when delivered to adolescents in a classroom environment. Results indicated that Bite Back did not lead to any significant improvement in mental health outcomes compared to the control condition. For the flourishing scale, there was a significant difference between the control group and Bite Back across the two time points, favoring control. However, the control condition started the intervention with lower scores and returned to the same level as the Bite Back condition post intervention, so it is difficult to establish if this was a case of regression to the mean or if there was a more positive impact of the control condition compared to Bite Back. The principle conclusion that can be drawn from these results is that there was no significant benefit of participating in the Bite Back condition compared to the control condition.

Previous findings from this team in a community sample of young people suggests that when Bite Back is delivered in an unstructured format and with freedom to choose engagement frequency, it led to reductions in mental health symptoms and improvements in well-being compared to the control condition [9,10]. The sample's age was comparable to our study sample as well as the baseline levels of psychopathology and well-being. However, our findings indicate that when the same program is delivered in a structured method and participants are obliged to use the website at a prescribed frequency, there is no benefit from the Bite Back. A plausible explanation for the current

XSL•FO
RenderX

findings is that it is the conversion of Bite Back into a structured school program that explains the observed results.

There are several aspects of the conversion that may have contributed to this. In the current study, participants were required to use Bite Back several times each week. Some students may have resented the pressure to interact with the program on a regular basis. Previous research has found that the benefits of acts of kindness and expressing gratitude disappeared when participants were asked to complete them more than once per week [30]. The authors of this study suggested that dose may be important when the participants were not free to choose the specific activity [31]. Other research has found that in an online positive psychology intervention completing 2-4 exercises improved well-being but attempting to complete up to 6 exercises negated all benefits [32]. These authors concluded that participants should not be overwhelmed with activities. Our findings suggest that in addition to these factors, being forced to engage in positive psychology may also remove its beneficial effects.

It has been previously found that individual factors such as motivation and effortful engagement are important in order to obtain benefits from positive psychology [33-35]. Level of engagement and choice of activity in the community study were up to participants, whereas in the current study, students were required to complete the exercises as part of their curricular activities at school. The greater level of motivation that can be expected when participants are completely free to choose their level of engagement in the program may have played a role in our findings.

Finally, the differential usage of specific aspects of Bite Back may have affected these findings. While the Bite Back program involves several activities relating to numerous positive psychology domains, there have been mixed findings in relation to the effectiveness of domain-specific well-being programs. Mindfulness, for example, has a wealth of research supporting its effectiveness as an intervention strategy when used in isolation [36,37]. However, the domain of optimism when used in isolation has significantly fewer supporting studies to date. Participants in the community study with free access to Bite Back content may have engaged to a greater extent with more “active ingredients”, such as mindfulness or with activities that piqued their interest, while participants in this study were required to engage with each positive psychology domain regardless of their interest. This may have dampened the treatment effects for the current study.

Our study highlights that the method of delivery of youth mental health programs may be equally important as content. Future research would also be warranted to clarify which delivery-specific factors may have been implicated in our study. While a number of such factors were considered in relation to these findings, only future research could determine whether these were indeed implicated and to what extent. It is possible that there were other delivery-specific factors not considered by us. Such research would benefit other school programs, which could then be tailored to accommodate these influences and also help transition programs from open-access to becoming more structured and school-based. In addition, while it appears plausible that delivery-specific factors account for our findings, it cannot be excluded that this study is a true reflection of the efficacy of Bite Back. Future trials to clarify this question would also be beneficial.

Limitations

The findings of this study should be interpreted in light of its limitations. The sample used in this study was a relatively heterogeneous sample with participants coming from schools that were both co-ed and single gender and that reflected different religious denominations and philosophies. There were relatively few checks on how the program was presented across the various schools and classrooms. This could be both a strength and limitation of the study, as it more closely replicated real-life applications of positive psychology in the school environment.

The sample size was relatively small, although this is usual in feasibility type research as it provides data for further large-scale studies. The study was also limited by a lack of long-term follow-up data, although it is common for the strongest reduction in symptoms to be observed post intervention and so it is likely that significant differences would not have appeared at a later time point.

Conclusion

The application of the online positive psychology program, Bite Back, into a structured classroom intervention did not demonstrate significant benefits when compared to a control intervention. This may not be a true reflection of the efficacy of the program but rather a problem in the application of Bite Back to the school environment.

Acknowledgments

This study was conducted with the support of a research grant from the Department of Health and Ageing. This work was completed at the Black Dog Institute at the University of New South Wales. We are also grateful to Professor Felicia Huppert for advice on the methodology to this study and to Jacqui Wallace and Nic Newling for technical assistance in running this trial. PB is supported by National Health and Medical Research Council fellowship 1035262.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.2 [17].

[PDF File (Adobe PDF File), 1MB - jmir_v17i7e187_app1.pdf]

References

1. Australian BO. National Survey of Mental Health and Wellbeing: Summary of Results. 2007. URL: http://www.abs.gov.au/ausstats/abs@.nsf/mf/4326.0 [accessed 2015-07-07] [WebCite Cache ID 6ZpxGGwvT]
2. Begg S, Vos T, Barker B, Stevenson C, Stanley L, Lopez A. The Burden of Disease and Injury in Australia. Australia: Institute of Health and Welfare; 2007. URL: http://www.aihw.gov.au/publication-detail/?id=6442467990 [accessed 2015-07-07] [WebCite Cache ID 6ZpxRLBS7]
3. Australian Bureau of Statistics. Mental Health of Young People. 2010. URL: http://www.abs.gov.au/ausstats/abs@.nsf/mf/4840.0.55.001/ [accessed 2015-07-07] [WebCite Cache ID 6ZpxWsP89]
4. Garnefski N, Legerstee J, Kraaij VV, Van Den Kommer T, Teerds J. Cognitive coping strategies and symptoms of depression and anxiety: a comparison between adolescents and adults. J Adolesc 2002 Dec;25(6):603-611. [Medline: 12490178]
5. Calear AL, Christensen H. Systematic review of school-based prevention and early intervention programs for depression. J Adolesc 2010 Jun;33(3):429-438. [doi: 10.1016/j.adolescence.2009.07.004] [Medline: 19647310]
6. Meyers J, Meyers B. Bi-directional influences between positive psychology and primary prevention. School Psychol Quart 2003;18(2):222.
7. Carvajal SC, Clair SD, Nash SG, Evans RI. Relating Optimism, Hope, and Self-Esteem to Social Influences in Deterring Substance Use in Adolescents. Journal of Social and Clinical Psychology 1998 Dec;17(4):443-465. [doi: 10.1521/jscp.1998.17.4.443]
8. Ciarrochi J, Heaven PC, Davies F. The impact of hope, self-esteem, and attributional style on adolescents' school grades and emotional well-being: A longitudinal study. Journal of Research in Personality 2007 Dec;41(6):1161-1178. [doi: 10.1016/j.jrp.2007.02.001]
9. Bite Back. URL: http://www.biteback.org.au/ [accessed 2014-02-21] [WebCite Cache ID 6NXtFqiDS]
10. Manicavasagar V, Horswood D, Burckhardt R, Lum A, Hadzi-Pavlovic D, Parker G. Feasibility and effectiveness of a web-based positive psychology program for youth mental health: randomized controlled trial. J Med Internet Res 2014;16(6):e140 [FREE Full text] [doi: 10.2196/jmir.3176] [Medline: 24901900]
11. Gomes AR, Marques B. Life skills in educational contexts: testing the effects of an intervention programme. Educational Studies 2013 May;39(2):156-166. [doi: 10.1080/03055698.2012.689813]
12. Suldo SM, Savage JA, Mercer SH. Increasing Middle School Students' Life Satisfaction: Efficacy of a Positive Psychology Group Intervention. J Happiness Stud 2013 Feb 1;15(1):19-42. [doi: 10.1007/s10902-013-9414-2]
13. Froh JJ, Sefick WJ, Emmons RA. Counting blessings in early adolescents: an experimental study of gratitude and subjective well-being. J Sch Psychol 2008 Apr;46(2):213-233. [doi: 10.1016/j.jsp.2007.03.005] [Medline: 19083358]
14. Marques SC, Lopez SJ, Pais-Ribeiro JL. "Building Hope for the Future": A Program to Foster Strengths in Middle-School Students. J Happiness Stud 2009 Dec 10;12(1):139-152. [doi: 10.1007/s10902-009-9180-3]
15. Proctor C, Tsukayama E, Wood AM, Maltby J, Eades JF, Linley PA. Strengths Gym: The impact of a character strengths-based intervention on the life satisfaction and well-being of adolescents. The Journal of Positive Psychology 2011 Sep 14;6(5):377-388. [doi: 10.1080/17439760.2011.594079]
16. Shoshani A, Steinmetz S. Positive Psychology at School: A School-Based Intervention to Promote Adolescents' Mental Health and Well-Being. J Happiness Stud 2013 Sep 24;15(6):1289-1311. [doi: 10.1007/s10902-013-9476-1]
17. 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]
18. Lyubomirsky S. The How of Happiness: A New Approach to Getting the Life You Want. New York: Penguin; 2007.
19. Lovibond S, Lovibond P. Manual for the Depression Anxiety Stress Scale. 2nd edition. Sydney: Psychology Foundation; 1995.
20. Brown TA, Chorpita BF, Korotitsch W, Barlow DH. Psychometric properties of the Depression Anxiety Stress Scales (DASS) in clinical samples. Behav Res Ther 1997 Jan;35(1):79-89. [Medline: 9009048]
21. Patrick J, Dyck M, Bramston P. Depression Anxiety Stress Scale: is it valid for children and adolescents? J Clin Psychol 2010 Sep;66(9):996-1007. [doi: 10.1002/jclp.20696] [Medline: 20694962]
22. Willemsen J, Markey S, Declercq F, Vanheule S. Negative emotionality in a large community sample of adolescents: the factor structure and measurement invariance of the short version of the depression anxiety stress scales (DASS-21). Stress and Health 2010 Jul 12;27(3):e120-e128. [doi: 10.1002/smi.1342]
23. Szabó M. The short version of the Depression Anxiety Stress Scales (DASS-21): factor structure in a young adolescent sample. J Adolesc 2010 Feb;33(1):1-8. [doi: 10.1016/j.adolescence.2009.05.014] [Medline: 19560196]
24. Huebner ES. Initial Development of the Student's Life Satisfaction Scale. School Psychology International 1991 Aug 01;12(3):231-240. [doi: 10.1177/0143034391123010]

XSL•FO
RenderX

25. Stewart-Brown S, Tennant A, Tennant R, Platt S, Parkinson J, Weich S. Internal construct validity of the Warwick-Edinburgh Mental Well-being Scale (WEMWBS): a Rasch analysis using data from the Scottish Health Education Population Survey. Health Qual Life Outcomes 2009;7:15 [FREE Full text] [doi: 10.1186/1477-7525-7-15] [Medline: 19228398]
26. Clarke A, Friede T, Putz R, Ashdown J, Martin S, Blake A, et al. Warwick-Edinburgh Mental Well-being Scale (WEMWBS): validated for teenage school students in England and Scotland. A mixed methods assessment. BMC Public Health 2011;11:487 [FREE Full text] [doi: 10.1186/1471-2458-11-487] [Medline: 21693055]
27. IBM Corp. SPSS Statistics for Windows. Version 22. Armonk, NY: IBM Corp; 2013.
28. Satterthwaite FE. An Approximate Distribution of Estimates of Variance Components. Biometrics Bulletin 1946 Dec;2(6):110. [doi: 10.2307/3002019]
29. Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale, NJ: L. Erlbaum Associates; 1988.
30. Lyubomirsky S, Sheldon KM, Schkade D. Pursuing happiness: The architecture of sustainable change. Review of General Psychology 2005;9(2):111-131. [doi: 10.1037/1089-2680.9.2.111]
31. Lyubomirsky S, Layous K. How Do Simple Positive Activities Increase Well-Being? Current Directions in Psychological Science 2013 Jan 22;22(1):57-62. [doi: 10.1177/0963721412469809]
32. Schueller SM, Parks AC. Disseminating self-help: positive psychology exercises in an online trial. J Med Internet Res 2012;14(3):e63 [FREE Full text] [doi: 10.2196/jmir.1850] [Medline: 22732765]
33. Layous K, Lee H, Choi I, Lyubomirsky S. Culture Matters When Designing a Successful Happiness-Increasing Activity: A Comparison of the United States and South Korea. Journal of Cross-Cultural Psychology 2013 May 09;44(8):1294-1303. [doi: 10.1177/0022022113487591]
34. Lyubomirsky S, Dickerhoof R, Boehm JK, Sheldon KM. Becoming happier takes both a will and a proper way: an experimental longitudinal intervention to boost well-being. Emotion 2011 Apr;11(2):391-402 [FREE Full text] [doi: 10.1037/a0022575] [Medline: 21500907]
35. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000 Jan;55(1):68-78. [Medline: 11392867]
36. Burke CA. Mindfulness-Based Approaches with Children and Adolescents: A Preliminary Review of Current Research in an Emergent Field. J Child Fam Stud 2009 Jun 27;19(2):133-144. [doi: 10.1007/s10826-009-9282-x]
37. Keng S, Smoski MJ, Robins CJ. Effects of mindfulness on psychological health: a review of empirical studies. Clin Psychol Rev 2011 Aug;31(6):1041-1056 [FREE Full text] [doi: 10.1016/j.cpr.2011.04.006] [Medline: 21802619]

Abbreviations

DASS-21: Depression, Anxiety, and Stress Scale—Short Form
SWEMWBS: Short Warwick Edinburgh Mental Well-Being Scale
SLSS: Student Life Satisfaction Scale

Edited by G Eysenbach; submitted 09.02.15; peer-reviewed by S Green, M Zhang, S Orlowski, D de Beurs; comments to author 24.06.15; revised version received 30.06.15; accepted 03.07.15; published 28.07.15.

Please cite as:
Burckhardt R, Manicavasagar V, Batterham PJ, Miller LM, Talbot E, Lum A
A Web-Based Adolescent Positive Psychology Program in Schools: Randomized Controlled Trial
J Med Internet Res 2015;17(7):e187
URL: http://www.jmir.org/2015/7/e187/
doi:10.2196/jmir.4329
PMID:26220564

©Rowan Burckhardt, Vijaya Manicavasagar, Philip J Batterham, Leonie M Miller, Elizabeth Talbot, Alistair Lum. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 28.07.2015. 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

Medium-Term Effectiveness of a Comprehensive Internet-Based and Patient-Specific Telerehabilitation Program With Text Messaging Support for Cardiac Patients: Randomized Controlled Trial

Ines Frederix[1,2], MD research doctorat; Dominique Hansen[1], PhD; Karin Coninx[3], PhD; Pieter Vandervoort[1,4], MD, PhD; Dominique Vandijck[1,5,6], PhD; Niel Hens[7,8], PhD (Statistics); Emeline Van Craenenbroeck[9], MD, PhD; Niels Van Driessche[1]; Paul Dendale[1,2], MD, PhD

[1]Mobile Health Institute, Faculty of Medicine & Life Sciences, Hasselt University, Diepenbeek, Belgium

[2]Cardiology, Department of Cardiology, Jessa Hospital, Hasselt, Belgium

[3]Expertise Center for Digital Media, Faculty of Sciences, Hasselt University, Hasselt, Belgium

[4]Cardiology, Department of Cardiology, Hospital East-Limburg, Genk, Belgium

[5]Patient Safety & Health Economics Institute, Department of Patient Safety & Health Economics, Hasselt University, Diepenbeek, Belgium

[6]Medicine and Health Sciences Institution, Department of Public Health, Ghent University, Ghent, Belgium

[7]Interuniversity Institute of Biostatistics and statistical Bioinformatics, Interuniversity Institute of Biostatistics and statistical Bioinformatics, Hasselt University, Diepenbeek, Belgium

[8]Vaccine and Infectious Disease Institute, Center for Health Economic Research and Modelling Infectious Diseases, University of Antwerp, Antwerp, Belgium

[9]Cardiology, Department of Cardiology, Antwerp University Hospital, Antwerp, Belgium

Corresponding Author:
Ines Frederix, MD research doctorat
Mobile Health Institute
Faculty of Medicine & Life Sciences
Hasselt University
Building D
Agoralaan
Diepenbeek, 3590
Belgium
Phone: 32 16583382
Fax: 32 16583382
Email: ines.frederix@gmail.com

Abstract

Background: Cardiac telerehabilitation has been introduced as an adjunct or alternative to conventional center-based cardiac rehabilitation to increase its long-term effectiveness. However, before large-scale implementation and reimbursement in current health care systems is possible, well-designed studies on the effectiveness of this new additional treatment strategy are needed.

Objective: The aim of this trial was to assess the medium-term effectiveness of an Internet-based, comprehensive, and patient-tailored telerehabilitation program with short message service (SMS) texting support for cardiac patients.

Methods: This multicenter randomized controlled trial consisted of 140 cardiac rehabilitation patients randomized (1:1) to a 24-week telerehabilitation program in combination with conventional cardiac rehabilitation (intervention group; n=70) or to conventional cardiac rehabilitation alone (control group; n=70). In the telerehabilitation program, initiated 6 weeks after the start of ambulatory rehabilitation, patients were stimulated to increase physical activity levels. Based on registered activity data, they received semiautomatic telecoaching via email and SMS text message encouraging them to gradually achieve predefined exercise training goals. Patient-specific dietary and/or smoking cessation advice was also provided as part of the telecoaching. The primary endpoint was peak aerobic capacity (VO_2 peak). Secondary endpoints included accelerometer-recorded daily step counts, self-assessed physical activities by International Physical Activity Questionnaire (IPAQ), and health-related quality of life (HRQL) assessed by the HeartQol questionnaire at baseline and at 6 and 24 weeks.

XSL•FO
RenderX

Results: Mean VO_2 peak increased significantly in intervention group patients (n=69) from baseline (mean 22.46, SD 0.78 mL/[min*kg]) to 24 weeks (mean 24.46, SD 1.00 mL/[min*kg], *P*<.01) versus control group patients (n=70), who did not change significantly (baseline: mean 22.72, SD 0.74 mL/[min*kg]; 24 weeks: mean 22.15, SD 0.77 mL/[min*kg], *P*=.09). Between-group analysis of aerobic capacity confirmed a significant difference between the intervention group and control group in favor of the intervention group (*P*<.001). At 24 weeks, self-reported physical activity improved more in the intervention group compared to the control group (*P*=.01) as did the global HRQL score (*P*=.01).

Conclusions: This study showed that an additional 6-month patient-specific, comprehensive telerehabilitation program can lead to a bigger improvement in both physical fitness (VO_2 peak) and associated HRQL compared to center-based cardiac rehabilitation alone. These results are supportive in view of possible future implementation in standard cardiac care.

(J Med Internet Res 2015;17(7):e185) doi:10.2196/jmir.4799

KEYWORDS

telemedicine; eHealth; effectiveness; Internet

Introduction

Cardiovascular disease (CVD) causes more than 4 million deaths in Europe and approximately 2 million deaths in the European Union each year, attributing to 47% and 40% of all deaths, respectively [1]. After an acute cardiovascular event, the European Society of Cardiology (ESC) guidelines recommend cardiac rehabilitation to prevent recurrent disease for both coronary artery disease (CAD) [2] and chronic heart failure (CHF) [3] patients (Class IB). However, the long-term benefits of conventional center-based cardiac rehabilitation are often disappointing because of lack of adherence to prescribed lifestyle behavior [4]. Therefore, it is important to examine and introduce adjunct intervention strategies to stimulate adherence to a healthy lifestyle.

During the past years, cardiac telerehabilitation was introduced as an adjunct or alternative to conventional cardiac rehabilitation to increase uptake rates, enable more prolonged care, and improve long-term success. Two recent systematic reviews concluded telerehabilitation to be noninferior and/or superior when compared to standard cardiac rehabilitation [5,6]. However, the European Heart Network emphasizes the need for more studies to be carried out on eHealth interventions to ensure its effectiveness and cost-effectiveness before large-scale implementation in current health care systems [7].

The aim of this multicenter, prospective randomized controlled trial was to assess medium-term effectiveness of a patient-specific, comprehensive cardiac telerehabilitation program in addition to standard ambulatory cardiac rehabilitation. Contrary to most prior clinical trials on cardiac telerehabilitation, it included both telemonitoring and telecoaching strategies and focused on multiple cardiac rehabilitation core components (physical activity, nutritional counseling, and smoking cessation) [5]. It was hypothesized that the addition of cardiac telerehabilitation to standard cardiac rehabilitation leads to significant greater increments in physical activity level and physical fitness. This paper reports on the main study results.

Methods

Patient Recruitment

Telerehab III (ISRCTN29243064) was a multicenter, prospective randomized controlled clinical trial run at Jessa Hospital (Hasselt) (n=103), Ziekenhuis-Oost Limburg (Genk) (n=27), and St Franciscus Hospital (Heusden-Zolder) (n=10) in Belgium between February 2013 and 2015. Patients were recruited/enrolled over a timeframe of 19 months (from February 2013 to August 2014). A detailed description of the study protocol has been published previously [8].

Patients were eligible for participation in Telerehab III when they entered cardiac rehabilitation for (1) CAD and treated conservatively with a percutaneous coronary intervention or with coronary artery bypass grafting, (2) CHF with reduced ejection fraction (EF; New York Heart Association [NYHA] classes I, II, and III), or (3) CHF with preserved EF (NYHA I, II, and III as defined in the ESC guidelines). Patients were required to have a computer at home with Internet access (they had to be computer and Internet literate). The main exclusion criteria were (1) CHF NYHA class IV, (2) symptomatic and/or exercise-induced cardiac arrhythmia within the previous 6 months, (3) physical disability related to musculoskeletal or neurological problems, and (4) severe cognitive impairment. All patients provided offline informed consent after the nature and possible consequences of the study were explained before study enrollment (see Multimedia Appendix 1 for the informed consent). Patients were recruited offline at the hospitals' rehabilitation centers by face-to-face information sessions. They were randomly assigned (1:1) to Internet-based telerehabilitation in addition to center-based rehabilitation (intervention group) or center-based rehabilitation alone (control group). A central computerized randomization system, using block randomization, ascertained equal distribution of patients in the different recruiting hospitals for both treatment arms.

The study was conducted in accordance with the principles stated in the Declaration of Helsinki (reviewed version of 2008), local and national regulations. The study protocol was approved by Jessa Ethics Committee (reference number: B243201216043). The trial is reported in accordance with CONSORT-EHEALTH (see Multimedia Appendix 2 for the completed CONSORT-EHEALTH form V1.6).

XSL•FO
RenderX

Study Intervention

Center-Based Cardiac Rehabilitation Program

Both groups participated in a 12-week conventional center-based cardiac rehabilitation program, including 45 pluridisciplinary rehabilitation sessions with at least 2 exercise training sessions per week [9]. Patients were instructed to exercise for 45 to 60 minutes per session at a target heart rate and/or workload corresponding to an intensity between their first ventilatory threshold (VT_1, as detected by V-slope method) and respiratory compensation point (RCP, as detected by carbon dioxide equivalent [VE/VCO_2] slope method). Endurance training consisted of walking/running and/or cycling and arm cranking. They also had at least one consultation with the dietician and the psychologist of the rehabilitation center. The dietician provided the patients with general guidelines on healthy diet, the psychologist aimed to improve the patient's self-efficacy to change prior unhealthy lifestyle behavior to a more healthy lifestyle behavior. He also assessed the patients' potential mood disorders (eg, depression, anxiety) related to their cardiac event.

Telerehabilitation Program

Intervention group patients received a 24-week, Internet-based, comprehensive telerehabilitation program in addition to the conventional center-based cardiac rehabilitation. The telerehabilitation program started at week 6 of the 12-week center-based cardiac rehabilitation allowing the intervention group patients to become familiarized with the telerehabilitation's motion sensor (Yorbody accelerometer, Belgium) and associated password-protected webservice during the 6-week overlap period. The program focused on multiple cardiac rehabilitation core components and used both physical activity telemonitoring and dietary/smoking cessation/physical activity telecoaching strategies. For the telemonitoring part, intervention group patients were prescribed patient-specific exercise training protocols based on achieved peak aerobic capacity (VO_2 peak) during initial maximal cardiopulmonary exercise testing and calculated body mass index (BMI) [8]. Intervention group patients were instructed to continuously wear the accelerometer and to regularly transmit their registered activity data to the telerehabilitation center's local server. They were instructed to transmit their physical activity data at least once weekly, but preferably daily. Data were transmitted to the telerehabilitation center's local server in a few minutes after starting transmission. These data enabled a semiautomatic telecoaching system to provide the patients with feedback via email and short message service (SMS) text messaging (once weekly), encouraging them to gradually achieve predefined exercise training goals (see Multimedia Appendix 3 for a screenshot of the website with an example SMS text message sent to the patient). In addition, patients received emails and/or SMS text messages (once weekly) with tailored dietary and smoking cessation recommendations. The dietary telecoaching program included a module for diabetes mellitus, arterial hypertension, obesity, and a healthy module. Cardiovascular risk factor profiling at entry of study determined which module(s) were prescribed for each patient. The smoking cessation telecoaching program included information on major risks associated with smoking, the health benefits of smoking cessation, and nicotine replacement therapy. It provided smokers with encouraging messages toward smoking cessation.

The content of the feedback messages differed from the content of the center-based cardiac rehabilitation program in that it changed over time based on how well the patient changed his prior lifestyle behavior. For example, the exercise training feedback was intended to encourage patients to achieve predefined patient-specific training goals. If a patient succeeded in getting closer toward these predefined goals, the feedback would encourage the patient to improve his/her training even more. If the patient's exercise training deteriorated during the study period, the feedback aimed to get the patient back on track. One independent person was responsible for technical assistance in case of sensor/system failure (part-time). One care provider supervised sent emails and/or SMS text messages and he/she was responsible for consistency and correctness of the content of sent messages. He/she also intervened in case of serious abnormal registrations (part-time). Access to registered data by the care provider was password-protected. The care provider that supervised sent emails and/or SMS text messages was a staff member that had coached cardiac patients for more than 5 years during their conventional center-based cardiac rehabilitation program. During the training period, this care provider also received a specific course on how to detect and what to do in case of alarming signs/symptoms. During the whole study period, one cardiologist supervised the care provider and was available to answer questions and to assist the care provider if necessary.

Outcome Measures

All outcome assessors were blinded to group allocation. The primary outcome measure was peak aerobic capacity (VO_2 peak); measured during maximal cardiopulmonary exercise testing [10] with breath-by-breath gas exchange analysis at baseline and after 6 and 24 weeks (Jaeger MS-CPX). The cardiopulmonary exercise test was maximal in case of an achieved heart rate >85% of the maximal predicted heart rate, a respiratory gas exchange ratio (RER) >1.1, and/or a ventilatory reserve (VR; VR=peak minute ventilation/maximal voluntary ventilation [VE peak/MVV]) >80% [10]. The first ventilatory threshold (VT_1) and the oxygen uptake efficiency slope (OUES) were used as surrogate markers for VO_2 peak in case of submaximal cardiopulmonary exercise test. VT_1 was defined by the V-slope method; OUES was calculated using the method of Baba et al [11]. Two independent investigators, blinded to treatment allocation, interpreted cardiopulmonary exercise test reports.

The first secondary outcome measure was daily physical activity [12], both registered by triaxial accelerometry (Yorbody sensor) and self-assessed by the patient. The accelerometer provided daily recordings of aerobic (defined as sustained activity at ≥60 steps/min for ≥10 minutes), regular (activity at <60 steps/min), and total (sum of aerobic and regular) steps. Self-reported physical activity was based on the offline International Physical Activity Questionnaire (IPAQ) questionnaire, completed at baseline and after 6 and 24 weeks. Metabolic equivalent task (MET) minutes were computed by multiplying predefined MET scores by the minutes of a specific activity performed to weigh

XSL•FO
RenderX

each type of activity by its energy requirement (for the domain leisure time activity and for all domains together). The following MET scores were used: 3.3 METs for walking, 4.0 METs for moderate physical activity, and 8.0 METs for vigorous physical activity.

Hemoglobin A_{1c} (HbA_{1c}), glycemic control, and lipid profile were assessed by blood sampling at study start and after 24 weeks study period.

The 14-item offline HeartQol questionnaire was used to assess health-related quality of life (HRQL) at study start and after 6 and 24 weeks [13]. Mean (SD) scores were calculated for both the physical (10-item) and emotional (4-item) subscale. The proportion of patients at the floor (*floor effect* defined as the lowest possible score on the questionnaire) and at the ceiling (*ceiling effect* defined as the best possible score) was determined to assess sensitivity to positive and negative changes in HRQL.

Qualitative feedback on the cardiac telerehabilitation system was obtained from intervention group patients by special offline feedback forms (see Multimedia Appendix 4 for an example of the feedback form used). Intervention group patients were requested to fill in these forms after study completion.

Statistical Analysis

Data analysis was performed using SPSS version 22 (SPSS Inc, Chicago, IL, USA) according to the intention-to-treat principle by assigned treatment group. Nonparametric alternatives were used for parametric statistics in case assumptions for the latter were violated. The Shapiro-Wilk test was used to assess normality. Paired *t* tests (parametric) or Wilcoxon signed rank tests (nonparametric) were used for within-group analysis; independent *t* tests (parametric) or Mann-Whitney *U* tests (nonparametric) for between-group analysis. Repeated measures ANOVA (parametric) or Friedman's ANOVA (nonparametric) compared multiple dependent means. Chi-square tests were used in case of categorical variables; Fisher's exact tests were used when expected frequencies were small. Pearson's (*r*) or Spearman's (ρ) correlation coefficients were calculated to express relationships between variables (bivariate correlations). The significance level for tests was 2-sided α=.05. Effect sizes for the HeartQol questionnaire were reported using the standardized response mean methodology (standardized response mean=[A–B]/D), where *A* and *B* are the mean scores at time 2 and time 1, respectively, and *D* represents the score change standard deviation [13]. Sensitivity analysis of accelerometric activity measurements was performed to cope with incomplete activity registrations. Inclusion thresholds of 1000, 2000, 3000, 4000, and 5000 total daily steps or ≥7, ≥8, and ≥9 daily measurement hours were arbitrarily chosen because these represented reliable registrations. All available data were used; no data imputation was performed for missing values. A priori sample size calculation yielded 140 necessary patients to detect a 20% effect size of the primary outcome measure (VO_2 peak) [14] between groups (intervention group vs control group) with a statistical power of 95% at a 2-sided type I error level of .05 and a dropout rate of 30%.

Results

A total of 140 patients agreed to participate in the study, 70 patients in the control group and 70 patients in the intervention group (Figure 1). The numbers and reasons for dropout during study period were similar for both treatment groups. Dropout patients were included in the final analysis, with the exception of one intervention patient who was diagnosed with a noncardiac-related pathology (ie, lung cancer) and was excluded from final analysis. Intervention patients transmitted their activity data a mean 1.0 (SD 0.3) times per week. When averaged over the whole study period (24 weeks), 76% (52/69) of the intervention group patients did more than 2000 total daily steps or measured their physical activities 8 hours or more per day. Both treatment groups had similar baseline demographics, clinical characteristics, and medication use (Table 1).

XSL•FO
RenderX

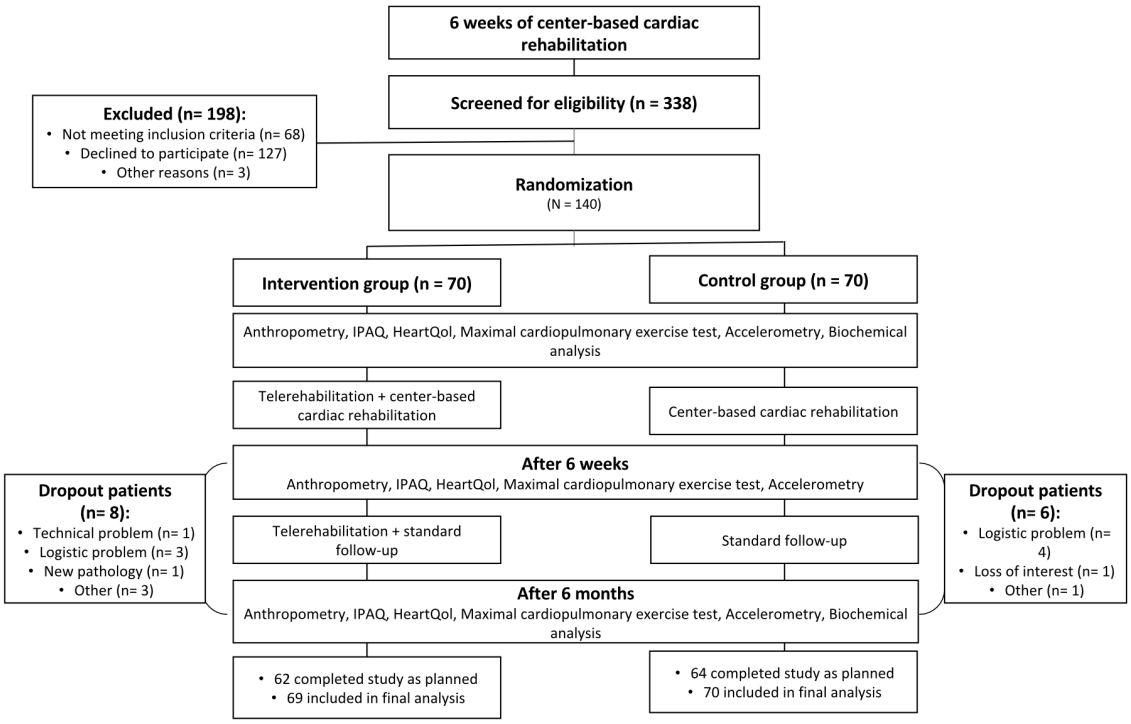
Table 1. Baseline demographics, clinical characteristics, and medication use (N=140).

Characteristic	Intervention group (n=69)	Control group (n=70)	*P*
Age (years), mean (SD)	61 (9)	61 (8)	.95
Gender, n (%)			.38
Female	10 (14)	15 (21)	
Male	59 (96)	55 (79)	
Type of cardiac pathology, n (%)			.53
Coronary artery disease	65 (94)	65 (93)	
Heart failure with reduced ejection fraction	2 (3)	4 (6)	
Heart failure with preserved ejection fraction	2 (3)	1 (1)	
NYHA class, [a] n (%)			.10
I	54 (78)	61 (87)	
II	12 (18)	4 (6)	
III	3 (4)	5 (7)	
Ejection fraction, n (%)			.32
>50%	52 (75)	50 (71)	
35%-50%	0 (0)	3 (4)	
<35%	17 (25)	17 (24)	
Comorbidity, n (%)			
Atrial fibrillation	5 (7)	6 (9)	.99
Diabetes mellitus	17 (25)	19 (27)	.85
Hyperlipidemia	53 (77)	55 (79)	.84
Arterial hypertension	40 (60)	44 (63)	.61
Family history of cardiac disease	34 (49)	36 (51)	.87
Peripheral artery disease	8 (12)	11 (16)	.62
Smoking, n (%)			.99
Current smoker	18 (26)	18 (26)	
Prior smoker	22 (32)	23 (33)	
Nonsmoker	29 (42)	29 (41)	
Body mass index (kg/m^2), mean (SD)	28 (5)	28 (4)	.54
Medications, n (%)			
On beta blocker	53 (77)	57 (81)	.61
On ACE-inhibitor	44 (64)	48 (69)	.72
On statin	66 (96)	64 (91)	.16
On antiplatelet therapy			.88
Dual antiplatelet therapy	37 (54)	40 (57)	
Antiplatelet monotherapy	29 (42)	27 (39)	
No antiplatelet therapy	3 (4)	3 (4)	
On diuretics	12 (17)	14 (20)	.76
On oral antidiabetics	10 (15)	10 (14)	.94
On insulin	7 (10)	5 (7)	.51
On anticoagulative therapy	4 (6)	5 (7)	.76
On antiarrhythmics	4 (6)	3 (4)	.67

XSL•FO
RenderX

[a]NYHA: New York Heart Association.

Figure 1. CONSORT patient flow diagram. IPAQ: International Physical Activity Questionnaire.



Aerobic Capacity

Table 2 shows the cardiopulmonary exercise test outcome measures assessed at baseline and after 6 and 24 weeks. Mean VO_2 peak improved significantly in intervention group patients from baseline (mean 22.46, SD 0.78 mL/[min*kg]) to 24 weeks (mean 24.46, SD 1.00 mL/[min*kg], P<.01). In the control group, mean VO_2 peak did not change after 24 weeks when compared to baseline (P=.09) and decreased from week 6 (mean 22.86, SD 0.66 mL/[min*kg]) to week 24 (mean 22.15, SD 0.77 mL/[min*kg], P=.02) after an initial nonsignificant increase. Between-group analysis of aerobic capacity was significant after 24 weeks (P<.001) in favor for the intervention group. VT_1 (Watt), OUES (mL/min/[log mL/min]), and Watt (% predicted) changed similarly over time (Figure 2) (see Multimedia Appendix 5 for a complete overview of the cardiopulmonary exercise test results).

XSL•FO
RenderX

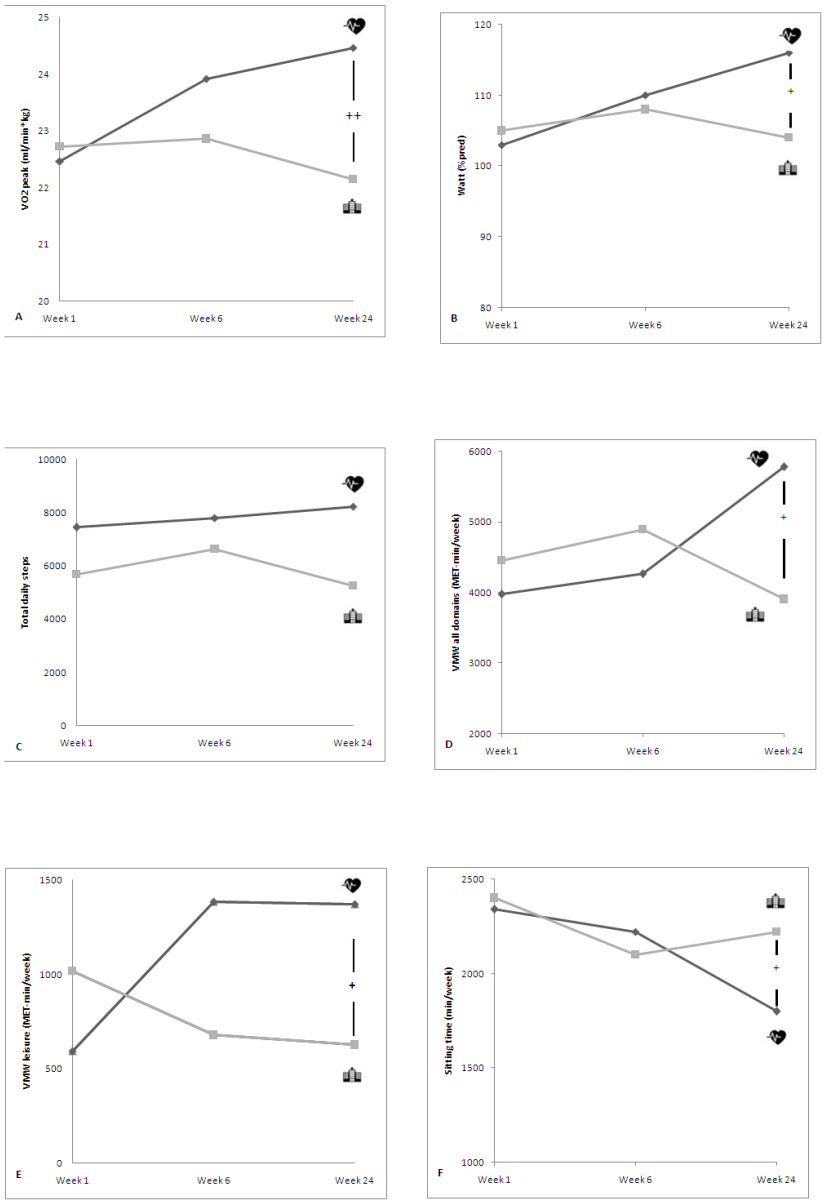
Table 2. Cardiopulmonary exercise test parameters at baseline and 6 weeks and 24 weeks follow-up period.

Cardiopulmonary exercise test results[a]	Week, mean (SD)			Within-group, *P*				Between-group, *P*			
	Week 1	Week 6	Week 24	Overall	Δ 6-1	Δ 24-6	Δ 24-1	Overall	Δ 6-1	Δ 24-6	Δ 24-1
Intervention group											
VO_2 peak (mL/[min*kg])	22.46 (6.43)	23.91 (6.74)	24.46 (7.57)	.01	.08	.38	.01	<.001	.19	.01	<.001
HR max (% pred)	79 (13)	80 (12)	83 (12)	.047	.99	.36	.05	.53	N/A	N/A	N/A
Watt (W)	152 (48)	163 (52)	165 (53)	.01	.02	.81	.01	.01	.90	.01	.02
Watt (pred%)	103 (23)	110 (27)	116 (27)	<.001	.01	.27	<.001	<.001	.83	.01	.01
VT_1 (W)	69 (24)	75 (25)	81 (26)	<.001	.74	<.001	<.001	<.001	.80	<.001	<.001
VT_1 (bpm)	93 (17)	91 (15)	96 (15)	.01	.99	.01	.08	.01	.35	.01	.047
OUES (mL/min/log[mL/min])	2067 (518)	2241 (545)	2272 (579)	<.001	.02	.045	<.001	.1	N/A	N/A	N/A
Weight (kg)	83.3 (18.2)	83.2 (17.4)	83.0 (17.3)	.69	N/A	N/A	N/A	.45	N/A	N/A	N/A
BMI (kg/m^2)	28 (5)	28 (5)	28 (5)	.63	N/A	N/A	N/A	.60	N/A	N/A	N/A
DBP rest (mm Hg)	82 (19)	81 (21)	77.24 (21.13)	.48	N/A	N/A	N/A	.67	N/A	N/A	N/A
SBP rest (mm Hg)	126 (21)	129 (30)	150 (140)	.26	N/A	N/A	N/A	.30	N/A	N/A	N/A
Control group											
VO_2 peak (mL/([min*kg])	22.72 (6.05)	22.86 (5.37)	22.15 (5.83)	.02	.99	.02	.09				
HR max (% pred)	77 (12)	79 (13)	79 (12)	.43	N/A	N/A	N/A				
Watt (W)	150 (49)	158 (50)	152 (53)	.01	<.001	.02	.99				
Watt (pred%)	105 (26)	108 (26)	104 (27)	.01	<.001	.01	.99				
VT_1 (W)	83 (34)	88 (34)	76 (31)	<.001	.20	<.001	.01				
VT_1 (bpm)	95 (15)	96 (15)	95 (17)	.53	N/A	N/A	N/A				
OUES (mL/min/log[mL/min])	2493 (2338)	2264 (637)	2142 (636)	.25	N/A	N/A	N/A				
Weight (kg)	82.7 (13.4)	82.5 (13.3)	82.5 (13.9)	.18	N/A	N/A	N/A				
BMI (kg/m^2)	28 (4)	28 (4)	27 (5)	.51	N/A	N/A	N/A				
DBP rest (mm Hg)	84 (21)	78 (19)	79 (17)	.33	N/A	N/A	N/A				
SBP rest (mm Hg)	129 (25)	127 (23)	129 (21)	.57	N/A	N/A	N/A				

[a] BMI: body mass index; CPET: cardiopulmonary exercise testing; DBP: diastolic blood pressure; HR: heart rate; N/A: not applicable; OUES: oxygen uptake efficiency slope; SBP: systolic blood pressure; VT_1: first ventilatory threshold.

XSL•FO
RenderX

Figure 2. Line charts depicting (A) mean VO2 peak (mL/[min*kg]), (B) mean Watt (% predicted), (C) median total daily steps, (D) median vigorous-moderate-walking (VMW) activity for all domains (MET-min/week), (E) median VMW activity for leisure time (MET-min/week), (F) median sitting time (min/week) for week 1, week 6, and week 24, respectively. Intervention group is represented by the heart icon; control group by the building icon. *P<.05, ** P<.001.



Physical Activity

Sensitivity analysis of accelerometric step data confirmed similar activity patterns for both groups, regardless of the thresholds (See Multimedia Appendix 6 for sensitivity analysis). More than 2000 total daily steps or 8 or more daily measurement hours were used as thresholds for further analysis. In the intervention group, total daily steps increased from baseline (median 7448, IQR 24) to both 6 weeks (median 7799, IQR 37) and 24 weeks (median 8233, IQR 32); however, no changes were significant (*P*=.24). In the control group, total daily steps showed an initial increasing trend from baseline (median 5678, IQR 13) to week 6 (median 6630, IQR 11), but declined afterwards (median 5265, IQR 17, *P*=.85). Total daily steps were positively correlated with VO_2 peak at baseline (ρ=.330, *P*=.01), 6 weeks (ρ=.237, *P*=.03), and 24 weeks (ρ=.485, *P*<.001).

Self-reported physical activity (by IPAQ questionnaire) was converted to MET-min/week of vigorous and/or moderate and/or walking (VMW) activities for the leisure time domain and all domains together, respectively. Summed leisure VMW increased significantly in the intervention group (based on Friedman's test, χ^2_2=13.7, *P*=.01) during the study period. In the control group, leisure VMW did not change (based on Friedman's test, χ^2_2=0.6, *P*=.72); however, it showed a downward trend (Figure 2). Between-group analysis confirmed a difference between the intervention group and control group in favor of the intervention group (*U*=1830, *z*=3.336, *P*=.01) (see Multimedia Appendix 7 for a summary of IPAQ questionnaire results). Contrary to the VMW activities, total sitting time decreased significantly during

XSL•FO
RenderX

the study period in the intervention group (based on Friedman's test, χ^2_2=19.9, P<.001). In the control group, total siting time did not change overall (based on Friedman's test, χ^2_2=3.7, P=.16). Control patients tended to decrease sitting time during the first 6 weeks, but increased their sitting time back again after 6 weeks (Figure 1). Between-group analysis confirmed a difference between the intervention group and control group for self-reported total sitting time (U=1360, z=–2.427, P=.02).

Cardiovascular Risk Factors

In the intervention group, no significant within-group differences were found for weight (P=.69), BMI (P=.63), diastolic blood pressure (P=.48), or systolic blood pressure (P=.26). The same was true for the control group (weight: P=.18; BMI: P=.51; diastolic blood pressure: P=.33; systolic blood pressure: P=.57). No between-group differences were found for these outcomes.

Fasting glucose levels, HbA_{1c}, and LDL cholesterol did not change during the study period in the intervention group (P=.67, P=.18, and P=.20, respectively) nor in the control group (P=.25, P=.51, and P=.31, respectively). Total cholesterol levels increased in both treatment groups, but no between-group differences were found (P=.97).

Health-Related Quality of Life

Intervention group patients showed a significant improvement in perceived HRQL for the physical subscale from baseline (mean 2.23, SD 0.08) to the end of study period (mean 2.52, SD 0.07; based on Friedman's test, χ^2_2=15.4, P<.001) (Table 3). The standardized response mean of 0.43 indicated a small to moderate effect size. Their global HRQL score also improved significantly (based on Friedman's test, χ^2_2=14.0, P<.001). The standardized response mean of 0.43 indicated a small to moderate effect size. The HRQL of the control group patients did not change during study period for the physical subscale (based on Friedman's test, χ^2_2=6.3, P=.05), the emotional subscale (based on Friedman's test, χ^2_2=0.5, P=.80), or the global scale (based on Friedman's test, χ^2_2=3.1, P=.21). Between-group analysis confirmed that globally the intervention group's HRQL improved more than the control group (U=2407, z=2.805, P=.01).

Qualitative Feedback

In all, 97% (67/69) of intervention group patients reported the telerehabilitation's motion sensor was easy to read and 97% (67/69) found it easy to use. In general, patients were very satisfied (44%, 30/69) or satisfied (51%, 35/69) (total: 95%, 65/69 very satisfied/satisfied) with the telerehabilitation program. In the end, 89% (61/69) of patients were willing to use the system after study completion.

XSL•FO
RenderX

Table 3. Results from HeartQol questionnaire at baseline and after 6 weeks and 24 weeks of follow-up.

Score	Intervention group[a]					Control group[a]					*P*
	Baseline	Week 6	Week 24	*P*	SRM	Baseline	Week 6	Week 24	*P*	SRM	
Physical subscale											
Mean (SD) score	2.23 (0.67)	2.45 (0.51)	2.52 (0.52)			2.27 (0.61)	2.39 (0.54)	2.28 (0.63)			
Ceiling effect (%)	12%	14%	24%			11%	16%	17%			
Floor effect (%)	0%	0%	0%			0%	0%	0%			
Overall				<.001	.44				.05	N/A	.01
Δ6-1				.046	.45				N/A	N/A	.05
Δ24-6				.99	N/A				N/A	N/A	.03
Δ24-1				<.001	.43				N/A	N/A	.01
Emotional subscale											
Mean (SD) score	2.36 (0.75)	2.47 (0.70)	2.53 (0.54)			2.41 (0.70)	2.43 (0.65)	2.41 (0.69)			
Ceiling effect (%)	30%	43%	34%			40%	35%	32%			
Floor effect (%)	5%	2%	0%			0%	0%	0%			
Overall				.14	N/A				.80	N/A	.22
Δ6-1				N/A	N/A				N/A	N/A	N/A
Δ24-6				N/A	N/A				N/A	N/A	N/A
Δ24-1				N/A	N/A				N/A	N/A	N/A
Global score											
Mean (SD) score	2.27 (0.63)	2.46 (0.51)	2.53 (0.44)			2.31 (0.59)	2.40 (0.51)	2.32 (0.58)			
Ceiling effect (%)	8%	13%	19%			6%	10%	10%			
Floor effect (%)	0%	0%	0%			0%	0%	0%			
Overall				.01	.44				.21	N/A	.01
Δ6-1				.07	.44				N/A	N/A	.05
Δ24-6				.84	N/A				N/A	N/A	.04
Δ24-1				<.001	.43				N/A	N/A	.01

[a]N/A: not applicable; SRM: standardized response mean.

Discussion

This study showed that an additional 6-month, patient-specific, comprehensive telerehabilitation program can lead to a bigger improvement in both physical fitness (VO_2 peak) and associated HRQL compared to center-based cardiac rehabilitation alone. The real difference between both groups occurred after center-based cardiac rehabilitation was completed. The VO_2 peak, daily total step count, and IPAQ's self-reported VMW activities increased from baseline to 6 weeks in both treatment groups. They additionally increased between weeks 6 and 24 in the intervention group, but decreased in the control group. Control group patients participated in the center-based cardiac rehabilitation during the first 6 weeks of the study period only, probably explaining their initial improvement in outcome measures. The observed findings imply that control group patients return to prior lifestyle behavior after center-based cardiac rehabilitation, whereas intervention group patients further maintain and ameliorate acquired behavioral change. The proportions of dropout patients in the recruiting hospitals (9%, 9/103 for Jessa Hospital; 7%, 2/27 for Ziekenhuis-Oost Limburg; and 30%, 3/10 for St. Franciscus Hospital) were relatively low compared with dropout rates in standard cardiac rehabilitation programs.

Recent literature findings confirmed telehealth interventions such as telemonitoring to be feasible and effective for heart failure patients [15-18]. Furthermore, 2 systematic reviews on cardiac teleinterventions were published [5,6] . We reported on cardiac telerehabilitation in CAD and CHF patients with a total of 13,248 patients enrolled in 37 studies and a mean follow-up

XSL•FO
RenderX

period of 9 months. We concluded that telerehabilitation was associated with significantly lower lack of adherence to physical activity guidelines (OR 0.56, 95% CI 0.45-0.69) [19-27]. However, Huang et al [6] found no statistically significant difference between telehealth interventions and center-based cardiac rehabilitation for exercise capacity (standardized mean difference [SMD]=−0.01, 95% CI −0.12 to 0.10), weight (SMD -0.13, 95% CI -0.30 to 0.05), systolic and diastolic blood pressure (mean difference [MD] -1.27, 95% CI -3.67 to 1.13 and MD 1.00, 95% CI -0.42 to 2.43, respectively), and lipid profile. Another recent systematic review by Widmer et al [28] on digital health interventions concluded that digital health interventions can improve cardiovascular risk factors such as weight loss, blood pressure, and LDL cholesterol in patients seeking primary prevention of CVD. In contrast, they found no consistent reductions in the aforementioned risk factors in secondary prevention studies.

The somewhat contrary findings between the review of Frederix et al [5], the results of the current Telerehab III trial (ISRCTN29243064) showing effectiveness of cardiac teleinterventions on exercise capacity, and the review of Huang et al [6] that showed no effect on exercise capacity, could be attributed to differences in intervention group programs. It appears that a comprehensive teleintervention, including at least physical activity telemonitoring and telecoaching, is necessary. The feedback provided by the teleintervention should be patient-specific to increase success rates.

In this Telerehab III trial, we found no significant effect of the additional cardiac telerehabilitation program on weight loss, blood pressure, lipid profile, and/or glycemic control. This is consistent with the findings of Huang et al [6] and Widmer et al [28]. Digital health interventions seem to be able to improve cardiovascular risk factors in primary prevention, but not secondary prevention programs. Future research should focus on furthering our understanding of the variables determining this success of digital health interventions in primary prevention populations, contrary to secondary prevention populations.

The intervention group patients could see and follow up their own transmitted activity data by logging onto the Telerehab III webpage as many times as they preferred. On average, they transmitted their activity data and logged onto the webpage a mean 1.0 (SD 0.3) times per week. For some patients, their frequency of data transmission increased during the study period; the frequency of others remained stable. There were almost no patients for whom the frequency of data transmission decreased during study period.

The reason for the increasing frequency of data transmission, seen for some of the intervention patients, remains unclear. In this trial, all intervention patients received feedback messages with the same frequency (once weekly). However, it would be interesting to investigate if the patients' frequency of data transmission would be different for different frequencies of sent feedback messages.

The strength of Telerehab III is that it, contrary to most analyzed trials in the review of Huang et al [6], provided intervention group patients with a comprehensive, patient-specific telerehabilitation program focusing on multiple core components (exercise training, nutritional counseling, smoking cessation). Both telecoaching and telemonitoring strategies were included; exercise training programs and dietary prescriptions were based on initial maximal cardiopulmonary exercise testing, BMI, and individual cardiovascular risk factor profile.

A limitation of this study was that Telerehab III was initially designed to recruit a broad cardiac patient population (including both CAD, CHF with reduced EF, and CHF with preserved EF). However, as shown by the baseline clinical characteristics (Table 1), only a minority of CHF patients eventually participated (5.8% and 7.1% in the intervention and control groups, respectively). This reduced the generalizability of study findings for CHF patients.

Sensitivity analysis of accelerometric activity measurements, based on arbitrarily selected thresholds, was performed to cope with incomplete activity registrations. Although the analysis found similar activity patterns for both groups regardless of chosen thresholds, it remains unclear whether the finally used thresholds of more than 2000 total daily steps or 8 or more daily measurement hours were most representative of reality. Therefore, one needs to interpret the accelerometric measurements with caution.

Finally, in Telerehab III, one part-time person (caregiver) was responsible for control of feedback content and one part-time person (technical assistant) for system/service operability. In a routine application setting, similar staff requirements would be sufficient.

This paper shows the addition of the cardiac telerehabilitation program to conventional center-based cardiac rehabilitation to be more effective than center-based cardiac rehabilitation alone in improving VO_2 peak, self-reported physical activity, and associated HRQL at 24 weeks. We plan to conduct a follow-up trial of Telerehab III (starting in August 2015) to assess whether the intervention-related health benefits persist 2 years after study termination. The current findings answer to the European Heart Network's question to profoundly well-document and evaluate critical eHealth interventions before large-scale deployment in the health care system. Future research should focus on even more elaborate comprehensive telerehabilitation programs that have the potential to improve not only aerobic capacity, physical activity level, and quality of life, but also improve the patient's cardiovascular risk factor profile (weight, blood pressure, lipids, and glycemia control).

Acknowledgments

The authors would like to thank Kim Bonne (physiotherapist), Toon Alders (physiotherapist), and Jan Berger (physiotherapist) for their support in patient recruitment. We would also like to acknowledge Evi Theunissen (study nurse) for her contribution to

XSL•FO
RenderX

patient inclusion and follow-up during study period and Steven Walscharts for his assistance in data interpretation and graphical designing of the figures. The Telerehab III trial was funded by a research grant from Flanders Care (grant number DEM2012-02-03) and from the Research Foundation Flanders (FWO; grant number: 1128915N). The authors/evaluators of the Telerehab III intervention are distinct from the intervention developers.

Authors' Contributions

IF participated in the design of the study, in the trial conduct, performed the statistical analysis, and drafted the manuscript. DH participated in the statistical analysis and helped to draft the manuscript. KC helped to draft the manuscript. PV helped to draft the manuscript. PD participated in the design of the study, in the statistical analysis, and helped to draft the manuscript. EVC, NH, DV, and NVD helped to draft the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Informed consent.

[PDF File (Adobe PDF File), 59KB - jmir_v17i7e185_app1.pdf]

Multimedia Appendix 2

CONSORT-EHEALTH checklist V1.6.2 [29].

[PDF File (Adobe PDF File), 150KB - jmir_v17i7e185_app2.pdf]

Multimedia Appendix 3

Screenshot of the website with an example SMS sent to the patient.

[PDF File (Adobe PDF File), 59KB - jmir_v17i7e185_app3.pdf]

Multimedia Appendix 4

Feedback form for Telerehab III intervention group.

[PDF File (Adobe PDF File), 11KB - jmir_v17i7e185_app4.pdf]

Multimedia Appendix 5

Complete set of measured cardiopulmonary exercise test parameters at baseline, 6 weeks and 24 weeks follow-up period.

[PDF File (Adobe PDF File), 103KB - jmir_v17i7e185_app5.pdf]

Multimedia Appendix 6

A Sensitivity analysis step data. B Use of telemonitoring system.

[PDF File (Adobe PDF File), 36KB - jmir_v17i7e185_app6.pdf]

Multimedia Appendix 7

Results from IPAQ questionnaire at baseline, 6 weeks and 24 weeks of follow-up period.

[PDF File (Adobe PDF File), 12KB - jmir_v17i7e185_app7.pdf]

References

1. Scharf P. European Heart Network: Fighting Heart Disease and Stroke. 2015. About us URL: http://www.ehnheart.org/about-us/overview.html [accessed 2015-05-31] [WebCite Cache ID 6YwSAyrVl]
2. Task FM, Montalescot G, Sechtem U, Achenbach S, Andreotti F, Arden C, ESC Committee for Practice Guidelines, et al. 2013 ESC guidelines on the management of stable coronary artery disease: the Task Force on the management of stable coronary artery disease of the European Society of Cardiology. Eur Heart J 2013 Oct;34(38):2949-3003 [FREE Full text] [doi: 10.1093/eurheartj/eht296] [Medline: 23996286]

XSL•FO
RenderX

3. McMurray JJ, Adamopoulos S, Anker SD, Auricchio A, Böhm M, Dickstein K, ESC Committee for Practice Guidelines. ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure 2012: The Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2012 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association (HFA) of the ESC. Eur Heart J 2012 Jul;33(14):1787-1847 [FREE Full text] [doi: 10.1093/eurheartj/ehs104] [Medline: 22611136]
4. Hansen D, Dendale P, Raskin A, Schoonis A, Berger J, Vlassak I, et al. Long-term effect of rehabilitation in coronary artery disease patients: randomized clinical trial of the impact of exercise volume. Clin Rehabil 2010 Apr;24(4):319-327. [doi: 10.1177/0269215509353262] [Medline: 20176771]
5. Frederix I, Vanhees L, Dendale P, Goetschalckx K. A review of telerehabilitation for cardiac patients. J Telemed Telecare 2015 Jan;21(1):45-53. [doi: 10.1177/1357633X14562732] [Medline: 25475219]
6. Huang K, Liu W, He D, Huang B, Xiao D, Peng Y, et al. Telehealth interventions versus center-based cardiac rehabilitation of coronary artery disease: A systematic review and meta-analysis. Eur J Prev Cardiol 2015 Aug;22(8):959-971. [doi: 10.1177/2047487314561168] [Medline: 25488550]
7. Scharf P. European Heart Network: Fighting Heart Disease and Stroke. 2015. EHN paper on eHealth URL: http://www.ehnheart.org/patients/papers/publication/698-ehn-paper-on-ehealth.html [accessed 2015-05-31] [WebCite Cache ID 6YwSKmss8]
8. Frederix I, Hansen D, Coninx K, Vandervoort P, Van Craenenbroeck EM, Vrints C, et al. Telerehab III: a multi-center randomized, controlled trial investigating the long-term effectiveness of a comprehensive cardiac telerehabilitation program - Rationale and study design. BMC Cardiovasc Disord 2015;15(1):29 [FREE Full text] [doi: 10.1186/s12872-015-0021-5] [Medline: 25948479]
9. Van Vlaenderen I. Cardiale revalidatie: klinische doeltreffendheid en gebruik in België. Brussels: Federaal Kenniscentrum voor de Gezondheidszorg (KCE); 2010. URL: https://kce.fgov.be/sites/default/files/page_documents/kce_140a_cardiale_revalidatie.pdf [accessed 2015-05-31] [WebCite Cache ID 6YwSh6Snn]
10. Ross RM. ATS/ACCP statement on cardiopulmonary exercise testing. Am J Respir Crit Care Med 2003 May 15;167(10):1451; author reply 1451. [doi: 10.1164/ajrccm.167.10.950] [Medline: 12738602]
11. Baba R, Nagashima M, Goto M, Nagano Y, Yokota M, Tauchi N, et al. Oxygen uptake efficiency slope: a new index of cardiorespiratory functional reserve derived from the relation between oxygen uptake and minute ventilation during incremental exercise. J Am Coll Cardiol 1996 Nov 15;28(6):1567-1572 [FREE Full text] [Medline: 8917273]
12. Warren JM, Ekelund U, Besson H, Mezzani A, Geladas N, Vanhees L, et al. Assessment of physical activity - a review of methodologies with reference to epidemiological research: a report of the exercise physiology section of the European Association of Cardiovascular Prevention and Rehabilitation. Eur J Cardiovasc Prev Rehabil 2010 Apr;17(2):127-139. [doi: 10.1097/HJR.0b013e32832ed875] [Medline: 20215971]
13. Oldridge N, Höfer S, McGee H, Conroy R, Doyle F, Saner H, (for the HeartQoL Project Investigators). The HeartQoL: part II. Validation of a new core health-related quality of life questionnaire for patients with ischemic heart disease. Eur J Prev Cardiol 2014 Jan;21(1):98-106. [doi: 10.1177/2047487312450545] [Medline: 22822180]
14. Frederix I, Van Driessche N, Hansen D, Berger J, Bonne K, Alders T, et al. Increasing the medium-term clinical benefits of hospital-based cardiac rehabilitation by physical activity telemonitoring in coronary artery disease patients. Eur J Prev Cardiol 2015 Feb;22(2):150-158. [doi: 10.1177/2047487313514018] [Medline: 24249840]
15. Agboola S, Jethwani K, Khateeb K, Moore S, Kvedar J. Heart failure remote monitoring: evidence from the retrospective evaluation of a real-world remote monitoring program. J Med Internet Res 2015;17(4):e101 [FREE Full text] [doi: 10.2196/jmir.4417] [Medline: 25903278]
16. Kitsiou S, Paré G, Jaana M. Effects of home telemonitoring interventions on patients with chronic heart failure: an overview of systematic reviews. J Med Internet Res 2015;17(3):e63 [FREE Full text] [doi: 10.2196/jmir.4174] [Medline: 25768664]
17. 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]
18. Zan S, Agboola S, Moore SA, Parks KA, Kvedar JC, Jethwani K. Patient engagement with a mobile web-based telemonitoring system for heart failure self-management: a pilot study. JMIR Mhealth Uhealth 2015;3(2):e33 [FREE Full text] [doi: 10.2196/mhealth.3789] [Medline: 25842282]
19. Barnason S, Zimmerman L, Nieveen J, Schulz P, Miller C, Hertzog M, et al. Influence of a symptom management telehealth intervention on older adults' early recovery outcomes after coronary artery bypass surgery. Heart Lung 2009;38(5):364-376 [FREE Full text] [doi: 10.1016/j.hrtlng.2009.01.005] [Medline: 19755186]
20. Hanssen TA, Nordrehaug JE, Eide GE, Hanestad BR. Does a telephone follow-up intervention for patients discharged with acute myocardial infarction have long-term effects on health-related quality of life? A randomised controlled trial. J Clin Nurs 2009 May;18(9):1334-1345. [doi: 10.1111/j.1365-2702.2008.02654.x] [Medline: 19220616]
21. Redfern J, Briffa T, Ellis E, Freedman SB. Choice of secondary prevention improves risk factors after acute coronary syndrome: 1-year follow-up of the CHOICE (Choice of Health Options In prevention of Cardiovascular Events) randomised controlled trial. Heart 2009 Mar;95(6):468-475. [doi: 10.1136/hrt.2008.150870] [Medline: 18801781]

XSL•FO
RenderX

22. Jelinek M, Vale MJ, Liew D, Grigg L, Dart A, Hare DL, et al. The COACH program produces sustained improvements in cardiovascular risk factors and adherence to recommended medications-two years follow-up. Heart Lung Circ 2009 Dec;18(6):388-392. [doi: 10.1016/j.hlc.2009.06.001] [Medline: 19648058]
23. de Lusignan S, Wells S, Johnson P, Meredith K, Leatham E. Compliance and effectiveness of 1 year's home telemonitoring. The report of a pilot study of patients with chronic heart failure. Eur J Heart Fail 2001 Dec;3(6):723-730. [Medline: 11738225]
24. Clark RA, Tideman P, Tirimacco R, Wanguhu K, Poulsen V, Simpson P, et al. A pilot study of the feasibility of an Internet-based electronic Outpatient Cardiac Rehabilitation (eOCR) program in rural primary care. Heart Lung Circ 2013 May;22(5):352-359. [doi: 10.1016/j.hlc.2012.11.005] [Medline: 23294762]
25. Piotrowicz E, Baranowski R, Bilinska M, Stepnowska M, Piotrowska M, Wójcik A, et al. A new model of home-based telemonitored cardiac rehabilitation in patients with heart failure: effectiveness, quality of life, and adherence. Eur J Heart Fail 2010 Feb;12(2):164-171. [doi: 10.1093/eurjhf/hfp181] [Medline: 20042423]
26. 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]
27. Kraal JJ, Peek N, Van den Akker-Van Marle ME, Kemps HM. Effects of home-based training with telemonitoring guidance in low to moderate risk patients entering cardiac rehabilitation: short-term results of the FIT@Home study. Eur J Prev Cardiol 2014 Nov;21(2 Suppl):26-31. [doi: 10.1177/2047487314552606] [Medline: 25354951]
28. Widmer RJ, Collins NM, Collins CS, West CP, Lerman LO, Lerman A. Digital health interventions for the prevention of cardiovascular disease: a systematic review and meta-analysis. Mayo Clin Proc 2015 Apr;90(4):469-480. [doi: 10.1016/j.mayocp.2014.12.026] [Medline: 25841251]
29. Eysenbach G, Consort- E. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

BMI: body mass index
CAD: coronary artery disease
CHF: chronic heart failure
EF: ejection fraction
ESC: European Society of Cardiology
HRQL: health-related quality of life
IPAQ: International Physical Activity Questionnaire
MET: metabolic equivalent task
NYHA: New York Heart Association
OUES: oxygen uptake efficiency slope
RCP: respiratory compensation point
SMS: short message service
VMW: vigorous-moderate-walking
VO2 peak: peak aerobic capacity

Edited by G Eysenbach; submitted 03.06.15; peer-reviewed by R Widmer; comments to author 03.07.15; accepted 07.07.15; published 23.07.15.

Please cite as:
Frederix I, Hansen D, Coninx K, Vandervoort P, Vandijck D, Hens N, Van Craenenbroeck E, Van Driessche N, Dendale P
Medium-Term Effectiveness of a Comprehensive Internet-Based and Patient-Specific Telerehabilitation Program With Text Messaging Support for Cardiac Patients: Randomized Controlled Trial
J Med Internet Res 2015;17(7):e185
URL: http://www.jmir.org/2015/7/e185/
doi:10.2196/jmir.4799
PMID:26206311

©Ines Frederix, Dominique Hansen, Karin Coninx, Pieter Vandervoort, Dominique Vandijck, Niel Hens, Emeline Van Craenenbroeck, Niels Van Driessche, Paul Dendale. Originally published in the Journal of Medical Internet Research

XSL•FO
RenderX

(http://www.jmir.org), 23.07.2015. 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

Adverse Drug Reaction Identification and Extraction in Social Media: A Scoping Review

Jérémy Lardon[1,2*], PhD; Redhouane Abdellaoui[3,4*], MS; Florelle Bellet[5], PharmD; Hadyl Asfari[2,6], PharmD; Julien Souvignet[2,6], MS; Nathalie Texier[4], PharmD; Marie-Christine Jaulent[6], PhD; Marie-Noëlle Beyens[7], MD; Anita Burgun[8,9], MD, PhD; Cédric Bousquet[2,6], PharmD, PhD

[1]Université Paris 13, Sorbonne Paris Cité, Laboratoire d'Informatique Médicale et d'Ingénieurie des Connaissances en e-Santé (LIMICS), (Unité Mixte de Recherche en Santé, UMR_S 1142), F-93430, Villetaneuse, France, Sorbonne Universités, University of Pierre and Marie Curie (UPMC) Université Paris 06, Unité Mixte de Recherche en Santé (UMR_S) 1142, Laboratoire d'Informatique Médicale et d'Ingénieurie des Connaissances en e-Santé (LIMICS), F-75006, Institut National de la Santé et de la Recherche Médicale (INSERM), U1142, Laboratoire d'Informatique Médicale et d'Ingénieurie des Connaissances en e-Santé (LIMICS), F-75006, Paris, France

[2]Service de Santé Publique et de l'Information Médicale (SSPIM), Department of Public Health and Medical Informatics, Centre Hospitalier Universitaire (CHU) University Hospital of Saint Etienne, Saint-Etienne, France

[3]Institut National de la Santé et de la Recherche Médicale (INSERM), Unité Mixte de Recherche en Santé (UMR_S) 1138, équipe 22, Centre de Recherche des Cordeliers, Université Paris Descartes, Sorbonne Paris Cité, F-75006, Paris, France

[4]Kappa Santé, Paris, France

[5]Centre de Pharmacovigilance, Centre Hospitalier Universitaire (CHU) University Hospital of Saint Etienne, Saint-Etienne, France

[6]Université Paris 13, Sorbonne Paris Cité, Laboratoire d'Informatique Médicale et d'Ingénieurie des Connaissances en e-Santé (LIMICS), (Unité Mixte de Recherche en Santé, UMR_S 1142), F-93430, Villetaneuse, France, Sorbonne Universités, University of Pierre and Marie Curie (UPMC) Université Paris 06, Unité Mixte de Recherche en Santé (UMR_S) 1142, Laboratoire d'Informatique Médicale et d'Ingénieurie des Connaissances en e-Santé (LIMICS), F-75006, Paris,, Institut National de la Santé et de la Recherche Médicale (INSERM), U1142, Laboratoire d'Informatique Médicale et d'Ingénieurie des Connaissances en e-Santé (LIMICS), F-75006, Paris, France

[7]Centre de Pharmacovigilance, Centre Hospitalier Universitaire (CHU) University Hospital of Saint Etienne, Saint-Etienne, France

[8]Institut National de la Santé et de la Recherche Médicale (INSERM), Unité Mixte de Recherche en Santé (UMR_S) 1138, équipe 22, Centre de Recherche des Cordeliers, Université Paris Descartes, Sorbonne Paris Cité, F-75006, Paris, France

[9]Assistance Publique-Hôpitaux de Paris (AP-HP), Hôpital Européen Georges-Pompidou (HEGP), Department of Medical Informatics, Paris, France

*these authors contributed equally

Corresponding Author:
Jérémy Lardon, PhD
Université Paris 13, Sorbonne Paris Cité, Laboratoire d'Informatique Médicale et d'Ingénieurie des Connaissances en e-Santé (LIMICS), (Unité Mixte de Recherche en Santé, UMR_S 1142), F-93430, Villetaneuse, France
Sorbonne Universités, University of Pierre and Marie Curie (UPMC) Université Paris 06, Unité Mixte de Recherche en Santé (UMR_S) 1142, Laboratoire d'Informatique Médicale et d'Ingénieurie des Connaissances en e-Santé (LIMICS), F-75006
Institut National de la Santé et de la Recherche Médicale (INSERM), U1142, Laboratoire d'Informatique Médicale et d'Ingénieurie des Connaissances en e-Santé (LIMICS), F-75006
Campus des Cordeliers, Institut National de la Santé et de la Recherche Médicale (INSERM) U 1142 - Laboratoire d'Informatique Médicale et d'Ingénieurie des Connaissances en e-Santé (LIMICS) Esc D - 2ème étage
15 rue de l'École de Médecine
Paris, 75006
France
Phone: 33 144279190
Fax: 33 144279192
Email: jeremy.lardon@chu-st-etienne.fr

Abstract

Background: The underreporting of adverse drug reactions (ADRs) through traditional reporting channels is a limitation in the efficiency of the current pharmacovigilance system. Patients' experiences with drugs that they report on social media represent a new source of data that may have some value in postmarketing safety surveillance.

Objective: A scoping review was undertaken to explore the breadth of evidence about the use of social media as a new source of knowledge for pharmacovigilance.

XSL•FO
RenderX

Methods: Daubt et al's recommendations for scoping reviews were followed. The research questions were as follows: How can social media be used as a data source for postmarketing drug surveillance? What are the available methods for extracting data? What are the different ways to use these data? We queried PubMed, Embase, and Google Scholar to extract relevant articles that were published before June 2014 and with no lower date limit. Two pairs of reviewers independently screened the selected studies and proposed two themes of review: manual ADR identification (theme 1) and automated ADR extraction from social media (theme 2). Descriptive characteristics were collected from the publications to create a database for themes 1 and 2.

Results: Of the 1032 citations from PubMed and Embase, 11 were relevant to the research question. An additional 13 citations were added after further research on the Internet and in reference lists. Themes 1 and 2 explored 11 and 13 articles, respectively. Ways of approaching the use of social media as a pharmacovigilance data source were identified.

Conclusions: This scoping review noted multiple methods for identifying target data, extracting them, and evaluating the quality of medical information from social media. It also showed some remaining gaps in the field. Studies related to the identification theme usually failed to accurately assess the completeness, quality, and reliability of the data that were analyzed from social media. Regarding extraction, no study proposed a generic approach to easily adding a new site or data source. Additional studies are required to precisely determine the role of social media in the pharmacovigilance system.

(J Med Internet Res 2015;17(7):e171) doi:10.2196/jmir.4304

KEYWORDS

pharmacovigilance; adverse drug reaction; Internet; Web 2.0; social media; text mining; scoping review; adverse event

Introduction

Pharmacovigilance is defined by the World Health Organization as "the science and activities relating to the detection, assessment, understanding and prevention of adverse effects or any other drug-related problem" [1]. It comprises postmarketing safety surveillance activities to monitor drug benefit/risk ratios and to identify new potential adverse drug reaction (ADR) signals in real-life conditions. An ADR is defined as "[...] a response to a medicinal product which is noxious and unintended and which occurs at doses normally used in man for the prophylaxis, diagnosis or therapy of disease or for the restoration, correction or modification of physiological function. (WHO, 1972)" [2]. Not all ADRs are identified during clinical trials because of their limited duration and the numbers and types of patients. ADRs need to be followed up after drug approval [3] and, therefore, are burdens for health systems that can potentially lead to hospitalization [4] or death [5].

Pharmacovigilance is mainly based on the spontaneous reporting of ADRs. Initially, only health professionals were allowed to report ADRs. Subsequently, however, a number of studies [6-10] demonstrated the value of patients as reporters. Hughes and Cohen stated that drug user reporting could be a complementary source of knowledge [11]. Currently, a number of countries consider direct patient reporting to be a valuable source in pharmacovigilance [12] and have implemented regulations and solutions for patients' spontaneous reporting to health authorities. Although patients have increased the number of reporters, ADR underreporting remains a limitation of the current pharmacovigilance system [13-15]. Moreover, other sources for pharmacovigilance are now considered, such as via the secondary use of electronic health records [16-23].

The recent Web 2.0 and social media expansions have been accompanied by a rapid growth in the number of discussions on the Internet regarding drug uses. Social media constitutes a new data source for postmarketing drug safety surveillance [24] and may be of interest in identifying signals because of their high volume and availability.

The use of Internet discussions as an additional data source relies on methods to parse, extract, structure, collect, and organize relevant information from the Web pages for analysis. The use of many sources, the large amount of data, and the heterogeneity of data require multiple steps to obtain analyzable corpora. Methods derived from big data and natural language processing (NLP) need to be considered. Recent works have proposed solutions to address these issues and to standardize the process of extracting information from Web pages in social media.

In addition, questions remain about the quality of information available in users' Web 2.0 discussions. Whereas electronic health records and health professionals' ADR reports are structured and well documented, there are no requirements regarding writing and structuring descriptions of pharmacovigilance-related events on social media, and information may be scarce or incomplete.

We performed a scoping review of relevant previously published studies to assess how social media can be used as a data source for postmarketing drug surveillance. This type of literature review aims at providing an overview of the type, extent, and quantity of research available on this topic. Our overview describes the methods used to manage the data from the corpus of Web users' messages and the obtained results, and identifies potential research gaps and future needs.

Methods

Overview

We used the scoping review methodology described by Arksey and O'Malley [25] and further refined by Levac et al [26] and Daudt et al [27]. This methodology divides reviews into six stages: (1) identifying the research question, (2) identifying relevant studies, (3) selecting studies, (4) charting the data, (5) collating, summarizing, and reporting the results, and (6)

XSL•FO
RenderX

consultation with stakeholders. Although the sixth stage is optional, we followed the recommendations to consider it a required component.

Stage 1: Identifying the Research Question

The focus of this scoping review is the use of social media as a new source of data in pharmacovigilance. We use the common definition of social media as media-based or user-generated content. Consequently, we did not consider the news media. To define the search question, we first selected a sample of available publications and found two types: (1) reviews of Web forums conducted by pharmacovigilance specialists on the one hand and (2) technical articles on information extraction approaches authored by computer scientists on the other. In this article, we use the term "identification" to denote the manual process of pulling up social media pages and reviewing them for reports of ADRs. The term "extraction" is used to describe the algorithms that automatically extract ADR information from social media. In the following sections, terms such as "messages," "social media," "discussion," and "page" refer to Web content.

The research question regarding the use of social media or pharmacovigilance is twofold:

1. Theme 1: What is the relevant information for ADR signals that have been issued from social media? The identification theme focuses on the first question and evaluates the information contained in patients' narrations on social media.

2. Theme 2: What are the methods used to extract information from social media? The extraction theme commits to describing the automated tools and methods that have been used to access structured and valuable pharmacovigilance information.

Stage 2: Identifying Relevant Studies

Two electronic databases—PubMed and Embase—were searched for English and French articles. The PubMed database was searched twice, as follows:

1. With the following keywords (query #1) to investigate the pharmacovigilance and social media dimensions: pharmacovigilance, adverse reaction, adverse event (AE), drug, medication, pharmaceutical product, social media, Web 2.0, social network, Twitter, Facebook, blog, forum, fora, message board, comment, and user feedback. An outline view of this request is presented in Figure 1.

2. Medical Subject Heading (MeSH) terms (query #2)—pharmacovigilance, natural language processing, Adverse Drug Reaction Reporting Systems, and Internet—associated with the following keywords in the title or the abstract: surveillance, Twitter, Facebook, Doctissimo (the main French health-related discussion forum), social media, social network, online health community, online discussion, medical data mining, online, patient forum, and natural language processing.

Query #3 was specially designed for the Embase database based on query #1. The details of the three queries are given in Table 1.

The upper date limit of June 2014 was applied, with no lower date limit, considering that articles published in the early days of social networks could be of interest.

As an iterative process and in accordance with the scoping review methodology, all references from the studies selected in stage 3 were screened, as were all of the publications that cited the selected studies. To broaden the scope of the search, Google Scholar was also used to search for citations.

XSL•FO
RenderX

Table 1. Full search strategy for each database.

Database	Query	Query text
PubMed		
	Query #1 (keywords)	(pharmacovigilance[MeSH[a] Terms] OR pharmacovigilance[All Fields] OR ADR[b][All Fields] OR ADE[c][All Fields] OR (("adverse reaction"[All Fields] OR "adverse event"[All Fields] OR "side effect"[All Fields]) AND (drug[All Fields] OR medication[All Fields] OR pharmaceutical product*[All Fields])) AND ("social media"*[All Fields] OR "Web 2.0"[TIAB[d]] OR "Web 2.0"[TIAB] OR "social media" [TIAB] OR "social network*" OR Twitter OR Facebook OR blog OR forum* OR fora OR message board* OR comment* OR (user feedback*))
	Query #2 (MeSH terms)	((((("pharmacovigilance"[MeSH]) OR surveillance[Title])) AND (((((Twitter[Title/Abstract]) OR Facebook[Title/Abstract]) OR Doctissimo[Title/Abstract])) OR (((((((((social media[Title/Abstract]) OR social networks[Title/Abstract]) OR "online health community"[Title/Abstract]) OR "online discussion"[Title/Abstract]) OR medical data mining[Title/Abstract]) OR online[Title/Abstract]) OR patient forum[Title/Abstract]) OR natural language processing[MeSH Terms]) OR "natural language processing"[Title/Abstract]))) OR (((((("Adverse Drug Reaction Reporting Systems"[MeSH]) AND (((((Twitter[Title/Abstract]) OR Facebook[Title/Abstract]) OR Doctissimo[Title/Abstract])) OR (((((((((social media[Title/Abstract]) OR social networks[Title/Abstract]) OR "online health community"[Title/Abstract]) OR "online discussion"[Title/Abstract]) OR medical data mining[Title/Abstract]) OR online[Title/Abstract]) OR patient forum[Title/Abstract]) OR natural language processing[MeSH Terms]) OR "natural language processing"[Title/Abstract])))) OR ((((((((((social media[Title/Abstract]) OR social networks[Title/Abstract]) OR "online health community"[Title/Abstract]) OR "online discussion"[Title/Abstract]) OR medical data mining[Title/Abstract]) OR online[Title/Abstract]) OR patient forum[Title/Abstract]) OR natural language processing[MeSH Terms]) OR "natural language processing"[Title/Abstract])) AND Adverse Drug Reaction Reporting Systems[MeSH Terms])) OR (("Adverse Drug Reaction Reporting Systems"[MeSH]) AND "Internet"[Mesh]))
Embase	Query #3	"pharmacovigilance"/de OR ADR OR ADE OR ("adverse reaction"/de OR "adverse event" OR "side effect"/de AND ("drug"/de OR "medication"/de OR "pharmaceutical product")) AND ("social media"/de OR "Web 2.0":ab,ti[e] OR "Web 2.0":ab,ti OR "social media":ab,ti OR "social network"/de OR Twitter OR Facebook OR blog OR forum OR fora OR "message board" OR comment OR "user feedback")

[a]MeSH: Medical Subject Heading

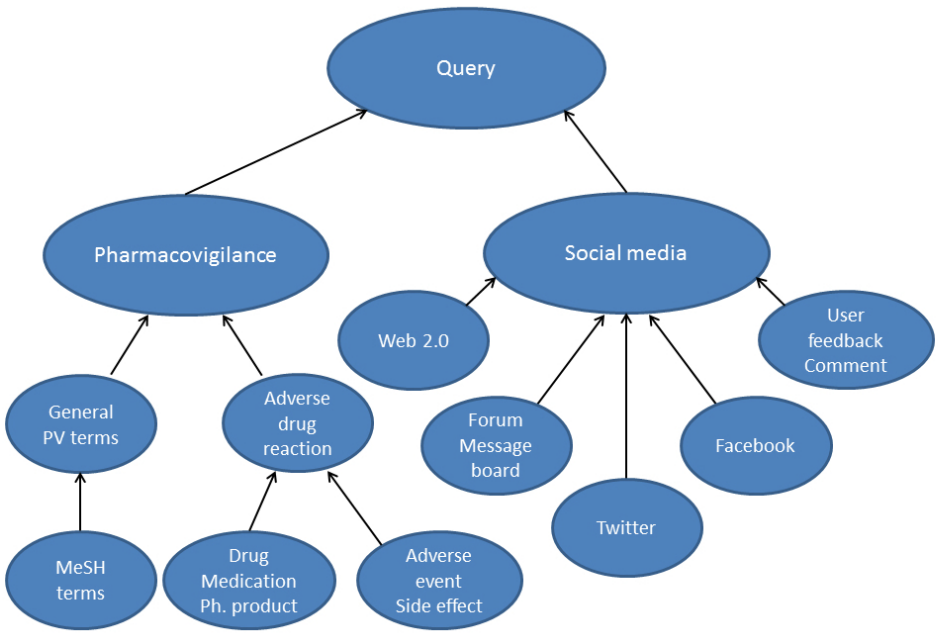
[b]ADR: adverse drug reaction

[c]ADE: adverse drug event

[d]TIAB: title and abstract

[e]ab, ti: abstract, title

Figure 1. Structure of the search queries.



Stage 3: Selecting Studies

Four authors (JL, RA, CB, and FB) independently screened the titles and abstracts (when available) of the query results to identify relevant articles. Disagreements about exclusions were discussed until a consensus was reached.

XSL•FO
RenderX

Abstracts were excluded if they met at least one of the following criteria:

1. Not related to drugs.

2. Not related to ADR reporting or ADR detection (eg, efficacy or effectiveness of a study's design).

3. Not related to patients' reporting (eg, a safety study in animals, signal detection on a pharmacovigilance database).

4. No, or insufficient, results on the use of social media.

5. The study was a review.

6. Soliciting reporting: the study used data from a source where patients were asked to report the ADRs. Patients' behavior is different depending on whether they are asked to report ADRs or they describe ADRs spontaneously without knowing that there may be further analysis of what they write [28].

7. Editorial: the study did not encompass a result but was an expression of the author's opinion about the usefulness of social media as a new source of knowledge for pharmacovigilance.

Stage 4: Charting the Data

Two pairs of reviewers independently identified a set of characteristics that could be used to describe the articles in each theme (FB and HA for theme 1, ie, identification; JL and RA for theme 2, ie, extraction). In addition to the basic elementary metadata, a number of characteristics were recorded for each theme. They are listed in Table 2.

The reviewers independently extracted data from the articles that were assigned to them.

Table 2. Article characteristics overview.

Characteristics	Theme 1	Theme 2
Year of publication	✓	✓
Language used in the studied texts	✓	✓
Type of data source, for example, forums or Twitter	✓	✓
Presence of an anonymization step	✓	✓
Volume of data analyzed	✓	✓
List of studied drugs	✓	✓
Coding ADRs[a] (medical lexicon)	✓	✓
Keywords the authors used to identify sources or posts of interest	✓	
Use of semiautomated processes (mixed methods)	✓	
Main results	✓	
Whether reported ADRs were highly informative or not	✓	
Seriousness of reported ADRs	✓	
Reference source was used for comparison with reported ADRs	✓	
Identification of potential unexpected ADRs or unexpected frequency of known ADRs	✓	
Analysis of the influence of other media, for example, television, radio, or the press, as a potential cause of increased ADR reporting in social media	✓	
If the authors mentioned the use of a crawler		✓
Implemented methods of preprocessing		✓
Lay language lexicon or tools used		✓
Authors attempted to identify the relationship between the drug and the event		✓
Authors used a machine-learning approach		✓
Evaluation of the extraction methods with metrics		✓
Comparison with external pharmacovigilance databases		✓
Whether the system enabled evaluating the unexpectedness of any extracted ADRs		✓

[a]ADR: adverse drug reactions

Stage 5: Collating, Summarizing, and Reporting Results

This work aimed to describe the methods and the results of the two themes. The results for theme 1, "identification," were related to studies based on the manual search and are presented in terms of methods and quality of data. The second theme, "extraction," was related to the studies that promoted an automated approach to extracting information from raw data. We summarized the "methods" sections of this last set of studies to describe each step of the methods presented.

XSL•FO
RenderX

Stage 6: Consultation

Following Daubt et al's recommendation [27], the research was multidisciplinary and multi-professional. The overall expertise covered pharmacovigilance, pharmacoepidemiology, public health, medical informatics, statistics, and data mining.

This helped us identify additional expectations regarding pharmacovigilance and social media, such as misuses, counterfeit drugs, drug-drug and food-drug interactions, and ADRs in specific populations such as pregnant women. It also permitted us to identify potential stakeholders—health care professionals, regulatory agencies, pharmaceutical companies, and patients—and establish the necessity of measuring the impact of mining social media, the interest in integrating this approach in a practical way in addition to classical reporting systems, and how we can be confident about the findings.

Results

Overview of Results

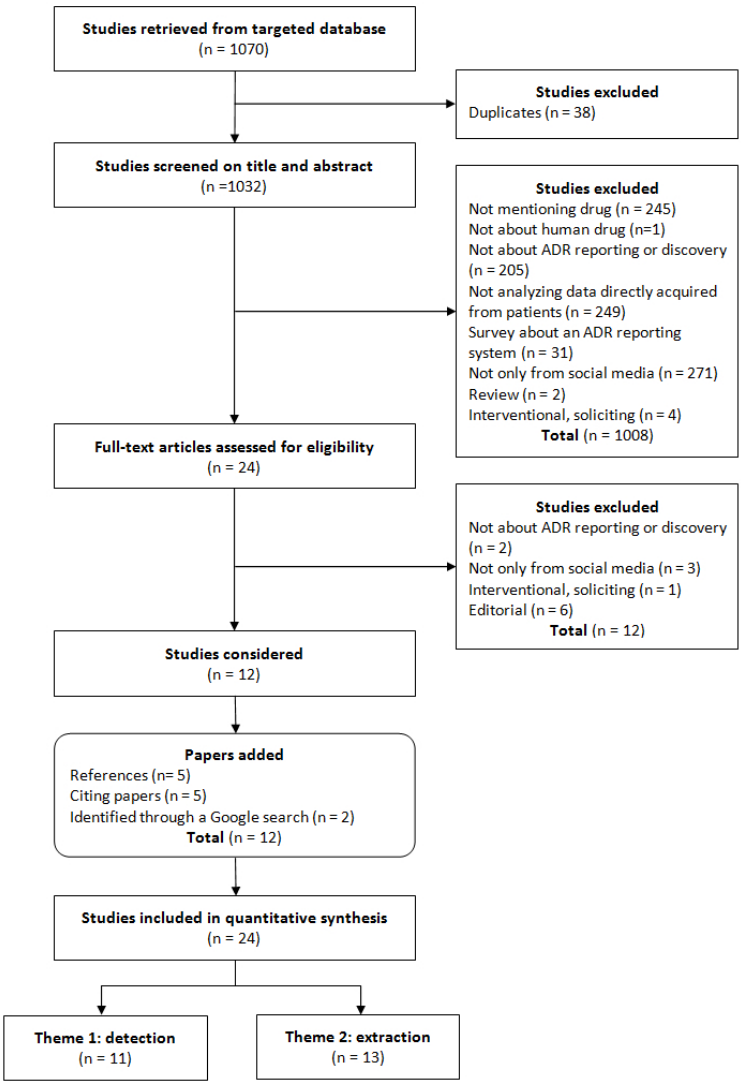
Figure 2 depicts the full review process and shows the number of citations excluded at each step.

A total of 1032 publications were identified in PubMed and Embase after duplicates (n=38) were removed. After applying the exclusion criteria to the titles, abstracts, and full texts, 11 citations were relevant to the research question at the end of the screening process (stage 3).

An additional 2 publications were added based on our personal previous knowledge, and 11 studies were identified by the references cited in the publications that were initially selected or by checking other articles that cited these publications on Google Scholar—7 via references and 4 via citing papers.

A total of 24 publications were finally included in the chart process. Of these, 11 (46%) were analyzed for theme 1 (identification) [11,28-37] and 13 (54%) for theme 2 (extraction) [38-50]. The detailed results of charting the data are displayed in Multimedia Appendix 1 and Multimedia Appendix 2.

Figure 2. Flowchart of our mapping process and study selection.



XSL•FO
RenderX

Theme 1: Identification

Overview

A total of 11 studies described a manual (or mixed) approach for identifying drug-ADR pairs in patients' narratives that were posted on social media. The majority of these studies were performed in the United States (6/11, 55%) or in France (3/11, 27%). Of these 11 studies, 4 (36%) were published in 2014 and 2 (18%) before 2010—in 2007 [36] and 2009 [34].

In 3 out of the 11 (27%) studies, the authors used the term "adverse event" rather than "adverse drug reaction" to refer to problems reported by patients in social media [30,32,35]. Pages et al justified this in their methodology by stating that the events reported by patients "were not analyzed by health professionals to assess the causal relationship" with the drug [35].

Table 3 details the steps identified in the studies to conduct the manual analysis: (1) selection of data sources, (2) data collection, (3) identification of drug-ADR/AE pairs, and (4) results evaluations.

Table 3. Main steps for identifying adverse drug reactions from social media.

Step	Description
Step 0: Selection of data sources	This step consists of identifying and selecting the most relevant websites to answer the research question. They can be identified using a combination of keywords (eg, generic or brand-name drug, disease, ADR[a]/AE[b]) in Web search engines.
Step 1: Data collection	Potentially relevant patient narratives or posts are identified by entering keywords into the search engine hosted by the selected websites (manual identification only) or using a semiautomated process. Data may be imported into software (after anonymization) with the aim of additional analyses.
Step 2: Identification of drug-ADR/AE pairs	The manual identification of drug-ADR/AE pairs is performed by reading the patients' narratives or posts that were initially collected.
Step 3: Results evaluation	This step consists of manually evaluating the frequency and the seriousness of the ADRs or AEs that were identified in patients' narratives or posts. The results can be compared, after coding, with those of other sources (Summary of Product Characteristics [SPC], clinical trials, pharmacovigilance databases, or literature) to identify potential new ADRs or an unexpected frequency of a known ADR.

[a]ADR: adverse drug reaction

[b]AE: adverse event

Analyzed Data Sources

The main data source was online forums. Three authors also reported on the analysis of Tweets or blogs [29,30,32]. The selected websites are often devoted to consumers' health. Patients' comments, mostly written in English, were identified by keywords (eg, brand-name and generic drugs, diseases, ADR) in the search engine hosted by the selected websites. In 2 of 11 (18%) studies, a hybrid (semiautomated) process was performed to identify potentially relevant posts [30,33].

The volume of analyzed data varied according to the studies—from 96 comments [31] to 61,401 Tweets [30]. Most often (ie, 8/11, 73%), the studies analyzed hundreds of narratives or posts.

Scope of the Surveillance

Out of 11 studies, 9 (82%) were designed to identify all of the ADRs/AEs that were potentially associated with one or more preselected drugs. Among these 9 studies, 7 (78%) focused on a class of drugs (eg, statins [31] or antineoplastic [33,35], psychotropic [10], or antiparkinsonian agents [36]), a recently marketed drug (eg, dabigatran [37]), or a drug that was removed from the market for pharmacovigilance reasons (eg, benfluorex [28]). Out of the 11 studies, 1 (9%) focused on two specific life-threatening ADRs—Stevens-Johnson syndrome and toxic epidermal necrolysis—and aimed at identifying any potentially associated drugs [29]. Another (1/11, 9%) was designed to identify and analyze predefined drug-AE pairs [34].

The rate of detected ADRs or AEs among the analyzed patients' comments varied according to the studies and was difficult to compare considering the methodological heterogeneity. Whereas Kmetz et al found a relatively low rate of reported AEs among patients' posts containing mentions of targeted drugs—0.3% of all brand mentions and 3.3% of brand mentions that contained side effects keywords [32]—Butt et al identified a large number of Internet descriptions for two rare and serious ADRs [29].

The informativeness of patients' comments was more or less evaluated in 8 of the 11 (73%) selected studies. In 5 of these 8 (63%) publications, information concerning patients' characteristics (ie, age, gender, and medical history), suspected drugs (ie, indications, dosages, and date of treatment initiation), or concomitant medications was often not available [31,32,34,35,37].

The presence of chronological criteria (ie, time to onset of ADRs, dechallenge, or rechallenge) was mentioned in only 3 of 11 (27%) studies and varied significantly according to the websites that were analyzed [10,31,36].

In 6 of 11 (55%) studies, the authors verified if the ADRs that were identified in social media were expected or not and compared them with those that had been reported in clinical trials (3/6, 50%) [33,36,37], in pharmacovigilance databases (2/6, 33%) [30,35], or in the studied drugs' Summary of Product Characteristics (SPC) (1/6, 17%) [31].

Out of 11 studies, 2 (18%) reported using a standard terminology—Medical Dictionary for Regulatory Activities (MedDRA) [30]—or Problem Intervention Documentation

XSL•FO
RenderX

(PI-Doc) [36], a German classification system, for coding the ADRs or AEs that were reported in social media. Freifeld et al [30] found that the AEs/ADRs identified in social media had similar profiles to those that were spontaneously reported through official channels, whereas Pages et al described the qualitative differences between the data sources [35].

Some studies identified potential unexpected ADRs [31,35,37] or unexpected frequencies of known ADRs [33,36]. Furthermore, Abou Taam et al retrospectively identified one case of severe valvulopathy 7 months before benfluorex was withdrawn because of this toxicity [28]. In their study, Butt et al compared patients' unsolicited Internet descriptions of severe cutaneous ADRs with experiences that had been previously collected in face-to-face interviews of survivors of these ADRs [29]. The authors identified new themes from Internet narratives, including fears and concerns of patients who had experienced the condition. According to the authors, patients also reported on more sensitive issues, such as sexual dysfunction, on the Internet rather than in face-to-face interviews.

The ADRs and AEs reported in social media were often less serious than those that were spontaneously reported through official channels [31,35], but they had impaired the patients' quality of life and their adherence to treatment [30,33,34]. Patients' comments were also related to complications from the ADRs, contraindications, drug-drug and food-drug interactions, and storage of drugs [37]. Users shared their experiences with individuals who were taking the same drug or had had similar adverse events, or with health care providers to obtain information or advice.

Abou Taam et al evaluated the impact of media coverage on benfluorex's withdrawal in France by analyzing patients' comments at three periods: before, during, and after the withdrawal of the drug. They found messages reflecting anxiety, anger, and other feelings, with drastic changes in consumers' perceptions following media coverage [28].

Theme 2: Extraction

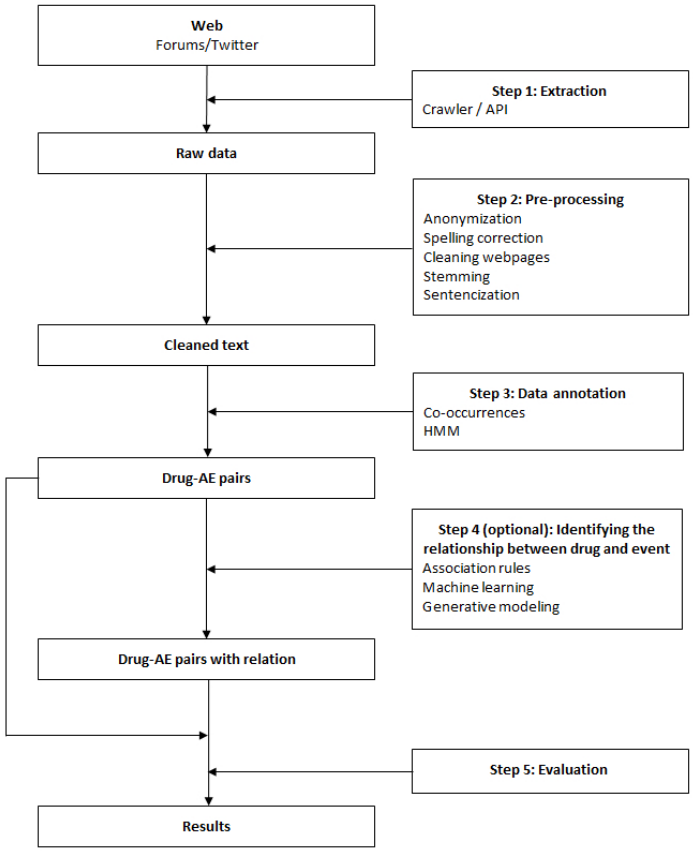
Overview

The 13 studies that were selected for the review had been recently published—2010 for the oldest [42].

Figure 3 shows a synthesis of the complete steps and presents five distinct parts: (1) data extraction, (2) preprocessing, (3) data annotation, (4) identifying the relationship between drug and event, and (5) results evaluation.

Multimedia Appendix 3 summarizes the use of these different steps in the papers.

Figure 3. Main steps for extraction of adverse drug reactions (ADRs) from social media.



Choice of the Source

The main data source was forum discussions in 12 studies out of 13 (92%) [34,35,37,38,40-44,47-49]. Out of the 13 studies, 1 (8%) [39] was about extracting narratives from Tweets.

Each study examined narratives that had been written in English, except for that of Hadzi-Puric and Grmusa [40]. The data volume was heterogeneous and varied from millions of messages [41] or billions of Tweets [39] to a more limited number of

XSL•FO
RenderX

messages, such as the 1290 messages included in the study by Hadzi-Puric and Grmusa [40].

The list of studied drugs was also heterogeneous. Out of the 13 selected studies, 11 (85%) focused on a limited number of drugs, such as lipid-modifying drugs in Li [43]. The other studies aimed at detecting signals of a large number of drugs, as in Liu and Chen's work [44], which considered all of the drugs from the Unified Modeling Language System (UMLS) and the US Food and Drug Administration's (FDA) Adverse Drug Event Reporting System (FAERS).

Data Extraction

The operating method to extract data from Web forums and social media depended on the nature of the source. For Web forums, 8 of 13 (62%) articles used an adapted Web crawler to collect Web pages, and then a Web scraper to extract the messages that were embedded in these Web pages [38,42-44,46-49].

Web scraping can be done through two approaches: (1) by taking the whole code of the page and cleaning it by eliminating the HTML tags and other unwanted elements or (2) by targeting the patients' messages using the HTML structure. The first approach was chosen by Benton [38]. In this work, approximately 48% of the tokens, defined as strings of characters delimited by whitespace in the original HTML pages, were retained to generate the corpus.

When the source was Twitter [39], specific application programming interfaces (APIs) were available for extracting data. These APIs provided some structured information, such as the date of the message or the pseudonym of the author, which are benefits to data quality, but the narratives still had to be processed with NLP.

Preprocessing Data

Using raw data extracted from social media or Web forums was not straightforward. "Preprocessing" the data was necessary and consisted, for example, of clarifying abbreviations and checking spelling mistakes.

As shown in Table 4, a number of types of transformations were performed on the extracted data.

Table 4. Transformations performed on the extracted data.

Transformation	Rationale and methods
Anonymization	Anonymization is required to remove patients' personal data to comply with medical confidentiality. Benton's team trained a classifier to determine if a token had to be anonymized or not [38]. Liu and Chen, only, did not extract the author pseudonyms [44], but they did not apply anonymization to the narratives.
Spelling correction	To maximize the detection of information in the corpus, spelling mistakes and typing errors that are common in texts extracted from social networks have to be corrected. The analyzed texts were extracted from social networks or public forums and included many abbreviations and typing errors. Li [43] applied this method to medical words that were often misspelled in messages.
Cleaning Web pages	Web pages consist of hundreds of tags that are invisible to users. When the crawler extracted a complete Web page code, a cleaning step was necessary to refine the content, as with Benton et al [38] and Liu and Chen [44].
Stemming	Reducing inflected words to their root helps to detect different forms of a word. This process reduces words to their word stem, base, or root forms, and these roots were then used for analysis. Different algorithms can be used by the «stemmer» [38,42,45,47,50]. For example, Benton et al [38] and Leaman et al [42] used the «Porter stemmer».
Sentencization/ Tokenization	Breaking the text up into segments of words, sentences, and paragraphs allows for analyzing the sentences and locutions in the corpus. Liu and Chen [44] used sentences at the information extraction level. Similarly, Benton et al [38] and Leaman et al [42] relied on a window of, respectively, 20 and 5 tokens in which the drug and the event co-occurred. Sentencization and tokenization are also documented in Liu and Chen [44], Nikfarjam and Gonzalez [45], and Yeleswarapu et al [50].

Annotation

Annotation of the corpus, for instance, identification of adverse events and drugs in messages, was performed in all of the studies reviewed in this theme. Annotation was realized by (1) machine-learning algorithms [39,44] and (2) final statistical evaluation [38-50].

Out of 13 studies, 9 (69%) used standard medical terminology, including Cerner Multum's Drug Lexicon, UMLS, side effect resource (SIDER), Coding Symbols for Thesaurus of Adverse Reaction Terms (COSTART), and MedDRA. Of the 13 studies, 8 (62%) took into account lay language. Among these 13 studies, 7 (54%) used lay vocabulary. Of the 13 that were originally selected, 3 (23%) studies [38,44,48] used a consumer health vocabulary [51], 1 of the 13 (8%) [49] used MedSyn [52], and 3 of the 13 (23%) [40,42,43] used a custom-built vocabulary. Of the 13 studies, 3 (23%) [39,40,44] mapped lay language to medical terminologies using MetaMap [53].

Relationships Between Drugs and Events

The relationship between the drug and the medical term was then analyzed. This relationship could have been an indication (ie, the drug was taken to treat the symptom or the disease), a cause (ie, the drug caused the pathology, in this case, an ADR), or a question about a potential causal relationship.

The methods were classified into two categories. The first category corresponds to methods that assessed a relationship between the medication and the event (ie, machine learning, association rules), which were used in 7 studies out of 13 (54%), with machine learning being used in 5 of 13 (38%) publications. When this approach was used, the evaluation was done thanks to cross-validation (3/13, 23%). The second category

XSL•FO
RenderX

corresponds to exploratory analysis to identify main safety themes from the corpus of messages (ie, statistically significant co-occurrences) [38,40].

Results Evaluation

In the studies that used a computerized approach, the evaluation of results was based on precision (9/13, 69%), recall (9/13, 69%), f-measure (6/13, 46%), accuracy (3/13, 23%), both true- and false-positive rates (1/13, 8%), log-likelihood ratios (1/13, 8%), support (1/13, 8%), confidence (1/13, 8%), leverage (1/13, 8%), and Bayesian confidence propagation neural network (BCPNN) scores and variance (1/13, 8%).

From the initial data volume, authors selected test sets on which they evaluated their systems. The sizes of the test samples were much smaller than the initial volumes of the extracted data. For example, Hadzi-Puric and Grmusa [40] and Li [43] used the whole initial data volume, whereas Yates [49] used only 480 posts out of the initial 400,000 extracted posts.

The pharmacovigilance database that was used for comparison was the FAERS [54] in 4 of the 13 (31%) studies. In 7 other studies out of 13 (54%), the annotators had varying expertise levels—from medical school students to pediatric clinicians and those with PhDs. Benton [38] referred to the tables and notes contained on the drug labels. Overall, only Li [43] did not document the constitution of a new gold standard or the use of an existing standard.

A majority of studies (7/13, 54%) were not only about expected ADRs, but also about discovering relationships between drugs and adverse events that had not been documented on the drug labels or in the literature.

Discussion

Gaps

This scoping review revealed some gaps among the selected studies that could be challenging to fill.

Although some studies that were related to the identification theme concluded that patients' comments posted in social media contained interesting data for pharmacovigilance (ie, potential unexpected ADRs, patients' risk perceptions, effects on adherence), they usually failed to accurately assess the completeness, quality, and reliability of these data. We could highlight the near absence of accessible information related to chronology (ie, time to onset, dechallenge, rechallenge) or differential diagnosis that would be necessary to assess the causal relationship between a drug and an AE. Moreover, evaluations of the seriousness and unexpectedness of ADRs was available in only a few studies. Finally, we retrieved no study that took into account exposures during pregnancy and only one study that partly focused on drug-drug or food-drug interactions.

More than the quality of the information shared in social media, issues can be raised about the reliability of this information. Indeed, social media users adopt pseudonyms, which may allow malicious persons to spread false rumors using multiple pseudonyms with limited risk of being identified as the origin of the rumor.

Furthermore, a user can post the same message twice or more on the same forum or on different forums using the same or different pseudonyms with no malevolent intent simply to maximize their chances of obtaining an answer. Consequently, it would be interesting to identify these duplicates. We found only one study in which an algorithm that addressed data redundancy was implemented but not described [40], and removing duplicates was seldom reported as an issue, for example, in Pages et al [35].

Regarding the extraction theme, we identified a set of processing steps that are used to process social media data after the Web crawling step and that could be recommended:

1. Anonymization: this was performed in only 2 studies out of 13 (15%), suggesting that privacy of data was not a major issue for the authors, who considered using pseudonyms to be sufficient for preserving confidentiality; nevertheless, it should be considered in every study that includes personal identifiers.

2. Preprocessing step: checking spelling errors and typographical errors; stemming, sentencization, and tokenization to process social media data.

3. Annotation and use of existing medical terminology.

Because none of the selected articles reported on a method that encompassed all of the steps we considered key, we assume that refining current methods and tools is desirable to improve the quality of processed data.

We also noticed that implementation did not follow a generic approach, which would be necessary for easily adding new sites or data sources. This is understandable in the context of a research project, but genericity should be addressed if more sites are intended to be included in the general pharmacovigilance process.

Finally, from the studies returned by our citation database queries, no study used comments on video-sharing websites as a source of data.

Limitations

The methodology has at least two limitations. First, when we constituted the research team, we were not exhaustive regarding the stakeholders we included in this review. For example, we lack stakeholders from regulatory agencies, from the pharmaceutical industry, and from patients or patient associations. The second limitation relates to the citation searches. We limited ourselves to PubMed and Embase. Although both of these resources offer a wide range of citations, we potentially missed some citations in the field, as illustrated by the fact that we selected two additional articles that were not found by the queries or by screening the citations. Moreover, the query itself was not trivial because the field is still a new research area. Finally, by using PubMed and Embase, we could not find any analyses of ADRs using social media that were conducted confidentially within company safety departments.

Perspectives

Emerging evidence on the effectiveness of social media for surveillance suggests that mining messages posted on social media may be helpful for complementing pharmacovigilance

XSL•FO
RenderX

systems. Examples of information retrieval from social media have previously been shown in other domains. For instance, it has been demonstrated that Tweets and restaurant reviews might aid in identifying and taking action on localized foodborne illnesses [55-57].

Adverse drug reactions are serious, underreported public health problems with high health and financial costs. A number of authors often described the cases of ADRs reported in social media as insufficiently informative to effectively assess a causal association with the drugs, compared with classical reporting in which quality criteria are available [58,59]. Moreover, extracting ADRs from social media presents specific technical constraints, given the unstructured information, compared with electronic health records or pharmacovigilance databases. Finally, spelling errors and patients' expressions [60,61] make extraction even more difficult.

However, our study confirms that there is a sufficient volume of data on pharmacovigilance in social media to work with and that quantity may eventually support the pharmacovigilance process despite variable quality. Whereas some websites may collect huge amounts of poorly documented posts, others such as PatientsLikeMe [60] collect very complete and high-quality data on drug treatments and, therefore, present very interesting possibilities for improving our knowledge on ADRs based on reliable information. It is thus necessary to further evaluate the quality of the different websites to fulfill the expectations of a new data source for pharmacovigilance.

Among the conceivable solutions for increasing reliability, we can suggest the use of comments' metadata (eg, pseudonym, date, and eventually the location given in the profile) to detect duplicated posts from the same author.

Indeed, the objective is to use social media as an additional source of data to expedite signals of potential ADRs. Local pharmacovigilance departments nationwide collect data on adverse events to track cases and interpret data for surveillance. Social media may help to detect the misuse or abuse (including overdose) of drugs [62,63] and adverse effects that would otherwise go unreported (eg, ADRs that are not serious but can impair the patients' quality of life and the adherence to treatment).

In order to verify the reliability of data retrieved online, comparison of this data with established sources, like FAERS or SIDER, as realized by several authors to derive reference material, can also be useful to detect new knowledge and improve quality of documentation of already described ADRs.

Social media may also provide new information on polypharmacy in real life, especially on the concomitant use of prescription drugs and self-medication drugs, and its consequences for patients, such as drug-drug interactions.

Nevertheless, it is necessary to verify how this new data source could be integrated into regular pharmacovigilance systems, with the aim of detecting, verifying, or validating signals.

Moreover, a number of authors highlighted the necessity of considering the context associated with the drug prescription, including whether any ADRs have been described in the media or discussed by regulatory agencies, to interpret the findings. Through the example of benfluorex, Abou Taam et al [28] analyzed narratives that were posted on French websites and reported drastic changes in consumers' risk perceptions following media coverage. As such, social media may be analyzed to assess consumers' behaviors and their risk perceptions and, finally, guide public communication campaigns.

Finally, a broader use of the Internet may include additional sources, such as soliciting reporting studies [64] we excluded, crowdsourcing [65] that may be complementary to social media, or Web search queries [66].

Conclusions

We conducted a scoping review to explore the potential interest in social media as a new source of data in pharmacovigilance and to define the methods for extracting data from this source. The exploratory aspect of the scoping review helped to give us an overview of this field, and this was a mandatory first step when we began our own work in the field. We are currently developing methods and tools within the Adverse Drug Reactions from Patient Reports In Social Media (ADR-PRISM) and *Vigilance dans les Forums sur le Médicament* (Vigi4MED) projects to collect data from social media and to evaluate the data's potential interest for pharmacovigilance. This scoping review was beneficial for identifying gaps in previous studies and designing our work plan.

Among the studies that were related to extraction, the oldest one was published in 2010, which shows that this field is still new and suggests that we can expect numerous further developments and improvements to come.

Finally, it appears that there are still outstanding questions about the data collected from social media and that there is sufficient room for improving extraction systems. Depending on the measured characteristics of social media as a new data source for pharmacovigilance and the headway in extracting ADRs, pharmacovigilants will have to define the role of social media in the classical pharmacovigilance system.

Acknowledgments

This work was funded by the grant AAP-2013-052 from the French agency for drug safety, *Agence Nationale de Sécurité du Médicament et des Produits de Santé* (ANSM), through the Vigi4MED project, and by the *Direction Générale des Entreprises* (DGE) and territorial collectivities (*Ile de France and Haute Normandie*) under the 16th *Fond Unique Interministériel* (FUI) request for proposals through the ADR-PRISM project.

XSL•FO
RenderX

Conflicts of Interest

Kappa Santé, the company that developed the Detec't tool, which extracts data from messages related to potential ADRs in social media, employs Redhouane Abdellaoui and Nathalie Texier. The other authors have no conflicts of interest with the subject matter discussed in the manuscript.

Multimedia Appendix 1

Characteristics of included studies from theme 1 (identification).

[XLSX File (Microsoft Excel File), 26KB - jmir_v17i7e171_app1.xlsx]

Multimedia Appendix 2

Characteristics of included studies from theme 2 (extraction).

[XLSX File (Microsoft Excel File), 17KB - jmir_v17i7e171_app2.xlsx]

Multimedia Appendix 3

Preprocessing transformations and analyses used in theme 2 (extraction) studies.

[XLSX File (Microsoft Excel File), 13KB - jmir_v17i7e171_app3.xlsx]

References

1. World Health Organization. Geneva, Switzerland: World Health Organization Pharmacovigilance URL: http://www.who.int/medicines/areas/quality_safety/safety_efficacy/pharmvigi/en/ [accessed 2015-06-23] [WebCite Cache ID 5zXGgxwBG]
2. World Health Organization. Geneva, Switzerland: World Health Organization English Glossary URL: http://www.who.int/medicines/areas/coordination/English_Glossary.pdf [accessed 2015-06-23] [WebCite Cache ID 6XyLFSz6M]
3. Classen DC, Resar R, Griffin F, Federico F, Frankel T, Kimmel N, et al. 'Global trigger tool' shows that adverse events in hospitals may be ten times greater than previously measured. Health Aff (Millwood) 2011 Apr;30(4):581-589 [FREE Full text] [doi: 10.1377/hlthaff.2011.0190] [Medline: 21471476]
4. Pirmohamed M, James S, Meakin S, Green C, Scott AK, Walley TJ, et al. Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. BMJ 2004 Jul 3;329(7456):15-19 [FREE Full text] [doi: 10.1136/bmj.329.7456.15] [Medline: 15231615]
5. Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. JAMA 1998 Apr 15;279(15):1200-1205. [Medline: 9555760]
6. Mitchell AS, Henry DA, Sanson-Fisher R, O'Connell DL. Patients as a direct source of information on adverse drug reactions. BMJ 1988 Oct 8;297(6653):891-893 [FREE Full text] [Medline: 3140967]
7. Medawar C, Herxheimer A. A comparison of adverse drug reaction reports from professionals and users, relating to risk of dependence and suicidal behaviour with paroxetine. Int J Risk Saf Med 2003;16(1):5-19.
8. van Grootheest K, de Jong-van den Berg L. Patients' role in reporting adverse drug reactions. Expert Opin Drug Saf 2004;3(4):363-368. [Medline: 15268652]
9. Blenkinsopp A, Wilkie P, Wang M, Routledge PA. Patient reporting of suspected adverse drug reactions: a review of published literature and international experience. Br J Clin Pharmacol 2007 Feb;63(2):148-156. [doi: 10.1111/j.1365-2125.2006.02746.x] [Medline: 17274788]
10. van Hunsel F, Talsma A, van Puijenbroek E, de Jong-van den Berg L, van Grootheest K. The proportion of patient reports of suspected ADRs to signal detection in the Netherlands: case-control study. Pharmacoepidemiol Drug Saf 2011 Mar;20(3):286-291. [doi: 10.1002/pds.2092] [Medline: 21351310]
11. Hughes S, Cohen D. Can online consumers contribute to drug knowledge? A mixed-methods comparison of consumer-generated and professionally controlled psychotropic medication information on the internet. J Med Internet Res 2011;13(3):e53 [FREE Full text] [doi: 10.2196/jmir.1716] [Medline: 21807607]
12. Margraff F, Bertram D. Adverse drug reaction reporting by patients: an overview of fifty countries. Drug Saf 2014 Jun;37(6):409-419. [doi: 10.1007/s40264-014-0162-y] [Medline: 24748428]
13. Bates DW, Evans RS, Murff H, Stetson PD, Pizziferri L, Hripcsak G. Detecting adverse events using information technology. J Am Med Inform Assoc 2003;10(2):115-128 [FREE Full text] [Medline: 12595401]
14. Tubert P, Bégaud B, Péré JC, Haramburu F, Lellouch J. Power and weakness of spontaneous reporting: a probabilistic approach. J Clin Epidemiol 1992 Mar;45(3):283-286. [Medline: 1569425]
15. Tubert P, Bégaud B. Random models for margins of a 2 x 2 contingency table and application to pharmacovigilance. Stat Med 1991 Jun;10(6):991-999. [Medline: 1876790]

XSL•FO
RenderX

16. Chen ES, Hripcsak G, Xu H, Markatou M, Friedman C. Automated acquisition of disease drug knowledge from biomedical and clinical documents: an initial study. J Am Med Inform Assoc 2008;15(1):87-98 [FREE Full text] [doi: 10.1197/jamia.M2401] [Medline: 17947625]
17. Trifirò G, Pariente A, Coloma PM, Kors JA, Polimeni G, Miremont-Salamé G, EU-ADR group. Data mining on electronic health record databases for signal detection in pharmacovigilance: which events to monitor? Pharmacoepidemiol Drug Saf 2009 Dec;18(12):1176-1184. [doi: 10.1002/pds.1836] [Medline: 19757412]
18. Aramaki E, Miura Y, Tonoike M, Ohkuma T, Masuichi H, Waki K, et al. Extraction of adverse drug effects from clinical records. Stud Health Technol Inform 2010;160(Pt 1):739-743. [Medline: 20841784]
19. Islamaj Doğan R, Névéol A, Lu Z. A context-blocks model for identifying clinical relationships in patient records. BMC Bioinformatics 2011;12 Suppl 3:S3 [FREE Full text] [doi: 10.1186/1471-2105-12-S3-S3] [Medline: 21658290]
20. Gurulingappa H, Mateen-Rajput A, Toldo L. Extraction of potential adverse drug events from medical case reports. J Biomed Semantics 2012;3(1):15 [FREE Full text] [doi: 10.1186/2041-1480-3-15] [Medline: 23256479]
21. Warrer P, Hansen EH, Juhl-Jensen L, Aagaard L. Using text-mining techniques in electronic patient records to identify ADRs from medicine use. Br J Clin Pharmacol 2012 May;73(5):674-684. [doi: 10.1111/j.1365-2125.2011.04153.x] [Medline: 22122057]
22. Eriksson R, Jensen PB, Frankild S, Jensen LJ, Brunak S. Dictionary construction and identification of possible adverse drug events in Danish clinical narrative text. J Am Med Inform Assoc 2013;20(5):947-953 [FREE Full text] [doi: 10.1136/amiajnl-2013-001708] [Medline: 23703825]
23. Li Y, Salmasian H, Vilar S, Chase H, Friedman C, Wei Y. A method for controlling complex confounding effects in the detection of adverse drug reactions using electronic health records. J Am Med Inform Assoc 2014;21(2):308-314 [FREE Full text] [doi: 10.1136/amiajnl-2013-001718] [Medline: 23907285]
24. Micoulaud-Franchi JA. [One step more toward pharmacovigilance 2.0. Integration of web data community for a pharmacovigilance more alert] [Article in French]. Presse Med 2011 Sep;40(9 Pt 1):790-792. [doi: 10.1016/j.lpm.2011.07.001] [Medline: 21802246]
25. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
26. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010;5:69 [FREE Full text] [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
27. Daudt HML, van Mossel C, Scott SJ. Enhancing the scoping study methodology: a large, inter-professional team's experience with Arksey and O'Malley's framework. BMC Med Res Methodol 2013;13:48 [FREE Full text] [doi: 10.1186/1471-2288-13-48] [Medline: 23522333]
28. Abou TM, Rossard C, Cantaloube L, Bouscaren N, Roche G, Pochard L, et al. Analysis of patients' narratives posted on social media websites on benfluorex's (Mediator®) withdrawal in France. J Clin Pharm Ther 2014 Feb;39(1):53-55. [doi: 10.1111/jcpt.12103] [Medline: 24304185]
29. Butt TF, Cox AR, Oyebode JR, Ferner RE. Internet accounts of serious adverse drug reactions: a study of experiences of Stevens-Johnson syndrome and toxic epidermal necrolysis. Drug Saf 2012 Dec 1;35(12):1159-1170. [doi: 10.2165/11631950-000000000-00000] [Medline: 23058037]
30. Freifeld CC, Brownstein JS, Menone CM, Bao W, Filice R, Kass-Hout T, et al. Digital drug safety surveillance: monitoring pharmaceutical products in twitter. Drug Saf 2014 May;37(5):343-350 [FREE Full text] [doi: 10.1007/s40264-014-0155-x] [Medline: 24777653]
31. Kheloufi F, Jean-Pastor MJ. Sharing experience with adverse drug reactions on internet: Which adverse drug reactions involving statins are described by patients on internet-based-forums? Fundam Clin Pharmacol 2014 May;28(Supplement s1):56 Abstracts of the 18th Annual Meeting of French Society of Pharmacology and Therapeutics, 81st Annual Meeting of Society of Physiology, 35th Pharmacovigilance Meeting, 15th APNET Seminar, 12th CHU CIC Meeting and 9th Annual Meeting of Physiology, Pharmacology and Therapeutics, , 22–24 April 2014, Poitiers, France [FREE Full text]
32. Kmetz J. Visible Technologies. 2013 Dec 26. Adverse event reporting: What pharmaceutical companies need to know URL: http://www.visibletechnologies.com/blog/adverse-event-reporting-pharma/ [accessed 2015-06-23] [WebCite Cache ID 6SftTDhuQ]
33. Mao JJ, Chung A, Benton A, Hill S, Ungar L, Leonard CE, et al. Online discussion of drug side effects and discontinuation among breast cancer survivors. Pharmacoepidemiol Drug Saf 2013 Mar;22(3):256-262 [FREE Full text] [doi: 10.1002/pds.3365] [Medline: 23322591]
34. Moncrieff J, Cohen D, Mason JP. The subjective experience of taking antipsychotic medication: a content analysis of Internet data. Acta Psychiatr Scand 2009 Aug;120(2):102-111. [doi: 10.1111/j.1600-0447.2009.01356.x] [Medline: 19222405]
35. Pages A, Bondon-Guitton E, Montastruc JL, Bagheri H. Undesirable effects related to oral antineoplastic drugs: comparison between patients' internet narratives and a national pharmacovigilance database. Drug Saf 2014 Aug;37(8):629-637. [doi: 10.1007/s40264-014-0203-6] [Medline: 25027671]
36. Schröder S, Zöllner YF, Schaefer M. Drug related problems with Antiparkinsonian agents: consumer Internet reports versus published data. Pharmacoepidemiol Drug Saf 2007 Oct;16(10):1161-1166. [doi: 10.1002/pds.1415] [Medline: 17486665]

XSL•FO
RenderX

37. Vaughan Sarrazin MS, Cram P, Mazur A, Ward M, Reisinger HS. Patient perspectives of dabigatran: analysis of online discussion forums. Patient 2014;7(1):47-54. [doi: 10.1007/s40271-013-0027-y] [Medline: 24030706]
38. Benton A, Ungar L, Hill S, Hennessy S, Mao J, Chung A, et al. Identifying potential adverse effects using the web: a new approach to medical hypothesis generation. J Biomed Inform 2011 Dec;44(6):989-996 [FREE Full text] [doi: 10.1016/j.jbi.2011.07.005] [Medline: 21820083]
39. Bian J, Topaloglu U, Yu F. Towards large-scale twitter mining for drug-related adverse events. In: Proceedings of the 2012 International Workshop on Smart Health and Wellbeing (SHB '12). New York, NY: ACM; 2012 Presented at: 2012 International Workshop on Smart Health and Wellbeing (SHB '12); October 29-November 2, 2012; Maui, HI. [doi: 10.1145/2389707.2389713]
40. Hadzi-Puric J, Grmusa J. Automatic drug adverse reaction discovery from parenting websites using disproportionality methods. In: Proceedings of the 2012 International Conference on Advances in Social Networks Analysis Mining (ASONAM 2012). Washington, DC: IEEE Computer Society; 2012 Presented at: 2012 International Conference on Advances in Social Networks Analysis Mining (ASONAM 2012); August 26-29, 2012; Istanbul, Turkey. [doi: 10.1109/ASONAM.2012.144]
41. Jiang Y, Liao QV, Cheng Q, Berlin RB, Schatz BR. Designing and evaluating a clustering system for organizing and integrating patient drug outcomes in personal health messages. AMIA Annu Symp Proc 2012;2012:417-426 [FREE Full text] [Medline: 23304312]
42. Leaman R, Wojtulewicz L, Sullivan R, Skariah A, Yang J, Gonzalez G. Towards internet-age pharmacovigilance: extracting adverse drug reactions from user posts to health-related social networks. In: Proceedings of the 2010 Workshop on Biomedical Natural Language Processing (BioNLP '10). Stroudsburg, PA: ACL; 2010 Presented at: 2010 Workshop on Biomedical Natural Language Processing (BioNLP '10); July 15, 2010; Uppsala, Sweden.
43. Li YA. DSpace@MIT. Cambridge, MA: Massachusetts Institute of Technology; 2011 Feb. Medical data mining: improving information accessibility using online patient drug reviews URL: http://dspace.mit.edu/handle/1721.1/66437 [accessed 2015-06-23] [WebCite Cache ID 6Sfu51j6d]
44. Liu X, Chen H. AZDrugMiner: An information extraction system for mining patient-reported adverse drug events in online patient forums. In: Proceedings of the 2013 International Conference on Smart Health (ICSH'13). Berlin, Heidelberg, Germany: Springer; 2013 Presented at: 2013 International Conference on Smart Health (ICSH'13); August 3-4, 2013; Beijing, China. [doi: 10.1007/978-3-642-39844-5_16]
45. Nikfarjam A, Gonzalez GH. Pattern mining for extraction of mentions of Adverse Drug Reactions from user comments. AMIA Annu Symp Proc 2011;2011:1019-1026 [FREE Full text] [Medline: 22195162]
46. Sampathkumar H, Chen XW, Luo B. Mining adverse drug reactions from online healthcare forums using hidden Markov model. BMC Med Inform Decis Mak 2014;14:91 [FREE Full text] [doi: 10.1186/1472-6947-14-91] [Medline: 25341686]
47. Wu H, Fang H, Stanhope SJ. Exploiting online discussions to discover unrecognized drug side effects. Methods Inf Med 2013;52(2):152-159. [doi: 10.3414/ME12-02-0004] [Medline: 23450374]
48. Yang CC, Jiang L, Yang H, Tang X. Detecting signals of adverse drug reactions from health consumer contributed content in social media. In: Proceedings of the ACM SIGKDD Workshop on Health Informatics. New York, NY: ACM; 2012 Presented at: ACM SIGKDD Workshop on Health Informatics; August 12, 2012; Beijing, China.
49. Yates A, Goharian N, Frieder O. Extracting adverse drug reactions from forum posts linking them to drugs. In: Proceedings of the 2013 ACM SIGIR Workshop on Health Search Discovery. New York, NY: ACM; 2013 Presented at: 2013 ACM SIGIR Workshop on Health Search Discovery; August 1, 2013; Dublin, Ireland. [doi: 10.1145/2484028.2484220]
50. Yeleswarapu S, Rao A, Joseph T, Saipradeep VG, Srinivasan R. A pipeline to extract drug-adverse event pairs from multiple data sources. BMC Med Inform Decis Mak 2014;14:13 [FREE Full text] [doi: 10.1186/1472-6947-14-13] [Medline: 24559132]
51. Consumer Health Vocabulary Initiative. URL: http://www.consumerhealthvocab.org/ [accessed 2015-06-23] [WebCite Cache ID 6Sfy1bTCY]
52. Yates A, Goharian N. ADRTrace: Detecting expected and unexpected adverse drug reactions from user reviews on social media sites. In: Proceedings of the 35th European Conference on Advances in Information Retrieval (ECIR'13). Berlin, Heidelberg, Germany: Springer-Verlag Berlin; 2013 Presented at: 35th European Conference on Advances in Information Retrieval (ECIR'13); March 24-27, 2013; Moscow, Russia p. 24-27. [doi: 10.1007/978-3-642-36973-5_92]
53. MetaMap. URL: http://metamap.nlm.nih.gov/ [accessed 2015-06-23] [WebCite Cache ID 6SfyBf7Xq]
54. US Food and Drug Administration. Silver Spring, MD: US FDA FDA Adverse Event Reporting System (FAERS) URL: http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Surveillance/AdverseDrugEffects/default.htm [accessed 2014-09-18] [WebCite Cache ID 6Sfxrss3N]
55. Harris JK, Mansour R, Choucair B, Olson J, Nissen C, Bhatt J, Centers for Disease Control and Prevention. Health department use of social media to identify foodborne illness - Chicago, Illinois, 2013-2014. MMWR Morb Mortal Wkly Rep 2014 Aug 15;63(32):681-685 [FREE Full text] [Medline: 25121710]
56. Kuehn BM. Agencies use social media to track foodborne illness. JAMA 2014 Jul;312(2):117-118. [doi: 10.1001/jama.2014.7731] [Medline: 24963655]
57. Chapman B, Raymond B, Powell D. Potential of social media as a tool to combat foodborne illness. Perspect Public Health 2014 Jul;134(4):225-230. [doi: 10.1177/1757913914538015] [Medline: 24990140]

XSL•FO
RenderX

58. Edwards IR, Lindquist M, Wiholm BE, Napke E. Quality criteria for early signals of possible adverse drug reactions. Lancet 1990 Jul 21;336(8708):156-158. [Medline: 1973481]
59. Franzen W. Can social media benefit drug safety? Drug Saf 2011 Sep 1;34(9):793. [doi: 10.2165/11595510-000000000-00000] [Medline: 21830842]
60. Smith CA, Wicks PJ. PatientsLikeMe: Consumer health vocabulary as a folksonomy. AMIA Annu Symp Proc 2008:682-686 [FREE Full text] [Medline: 18999004]
61. Zeng Q, Kogan S, Ash N, Greenes RA, Boxwala AA. Characteristics of consumer terminology for health information retrieval. Methods Inf Med 2002;41(4):289-298. [Medline: 12425240]
62. Hanson CL, Cannon B, Burton S, Giraud-Carrier C. An exploration of social circles and prescription drug abuse through Twitter. J Med Internet Res 2013;15(9):e189 [FREE Full text] [doi: 10.2196/jmir.2741] [Medline: 24014109]
63. Hanson CL, Burton SH, Giraud-Carrier C, West JH, Barnes MD, Hansen B. Tweaking and tweeting: exploring Twitter for nonmedical use of a psychostimulant drug (Adderall) among college students. J Med Internet Res 2013;15(4):e62 [FREE Full text] [doi: 10.2196/jmir.2503] [Medline: 23594933]
64. Härmark L, van Puijenbroek E, van Grootheest K. Intensive monitoring of duloxetine: results of a web-based intensive monitoring study. Eur J Clin Pharmacol 2013 Feb;69(2):209-215 [FREE Full text] [doi: 10.1007/s00228-012-1313-7] [Medline: 22688722]
65. Gottlieb A, Hoehndorf R, Dumontier M, Altman RB. Ranking adverse drug reactions with crowdsourcing. J Med Internet Res 2015;17(3):e80 [FREE Full text] [doi: 10.2196/jmir.3962] [Medline: 25800813]
66. 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]

Abbreviations

ab,ti: abstract, title
ADE: adverse drug event
ADR: adverse drug reaction
ADR-PRISM: Adverse Drug Reactions from Patient Reports In Social Media
AE: adverse event
ANSM: the French drug safety agency (French acronym of Agence Nationale de Sécurité du Médicament et des Produits de Santé)
AP-HP: Assistance Publique-Hôpitaux de Paris
API: application programming interface
BCPNN: Bayesian confidence propagation neural network
CHU: University Hospital Center (French acronym of Centre Hospitalier Universitaire)
COSTART: Coding Symbols for Thesaurus of Adverse Reaction Terms
DGE: block grants for investment expenditures (French acronym of Direction Générale des Entreprises)
FAERS: FDA Adverse Drug Event Reporting System
FDA: Food and Drug Administration
FUI: French acronym of Fond Unique Interministériel
HEGP: Hôpital Européen Georges-Pompidou
INSERM: the French National Institute for Health and Medical Research (French acronym of Institut National de la Santé et de la Recherche Médicale)
LIMICS: Laboratoire d'Informatique Médicale et d'Ingénieurie des Connaissances en e-Santé
MedDRA: Medical Dictionary for Regulatory Activities
MeSH: Medical Subject Heading
NLP: natural language processing
PI-Doc: Problem Intervention Documentation
SIDER: side effect resource
SPC: Summary of Product Characteristics
TIAB: title and abstract
SSPIM: Department of Public Health and Medical Informatics (French acronym of Service de Santé Publique et de l'Information Médicale)
UMLS: Unified Modeling Language System
UMR_S: Unité Mixte de Recherche en Santé
UPMC: University of Pierre and Marie Curie
Vigi4MED: French acronym of Vigilance dans les Forums sur le Médicament

XSL•FO
RenderX

Edited by G Eysenbach; submitted 30.01.15; peer-reviewed by N Dasgupta, M Chary; comments to author 27.02.15; revised version received 09.04.15; accepted 22.04.15; published 10.07.15.

Please cite as:
Lardon J, Abdellaoui R, Bellet F, Asfari H, Souvignet J, Texier N, Jaulent MC, Beyens MN, Burgun A, Bousquet C
Adverse Drug Reaction Identification and Extraction in Social Media: A Scoping Review
J Med Internet Res 2015;17(7):e171
URL: http://www.jmir.org/2015/7/e171/
doi:10.2196/jmir.4304
PMID:26163365

©Jérémy Lardon, Redhouane Abdellaoui, Florelle Bellet, Hadyl Asfari, Julien Souvignet, Nathalie Texier, Marie-Christine Jaulent, Marie-Noëlle Beyens, Anita Burgun, Cédric Bousquet. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 10.07.2015. 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

Overcoming Clinical Inertia: A Randomized Clinical Trial of a Telehealth Remote Monitoring Intervention Using Paired Glucose Testing in Adults With Type 2 Diabetes

Deborah A Greenwood[1,2], RN, PhD; Shelley A Blozis[3], PhD; Heather M Young[4], RN, PhD, FAAN; Thomas S Nesbitt[5], MD, MPH; Charlene C Quinn[6], RN, PhD, FAAN

[1]Clinical Performance Improvement Consultant, Office of Patient Experience, Quality and Clinical Effectiveness, Sutter Health, Sacramento, CA, United States

[2]Program Coordinator, Sutter Health Integrated Diabetes Education Network, Sutter Medical Foundation, Sacramento, CA, United States

[3]Associate Professor, Department of Psychology, University of California Davis, Davis, CA, United States

[4]Associate Vice Chancellor for Nursing, Dean and Professor, Betty Irene Moore School of Nursing, University of California Davis, Sacramento, CA, United States

[5]Associate Vice Chancellor, Strategic Technologies and Alliances, Director, Center for Health and Technology, University of California Health System, Sacramento, CA, United States

[6]Associate Professor, Department of Epidemiology and Public Health, School of Medicine, Univeristy of Maryland, Baltimore, MD, United States

Corresponding Author:
Deborah A Greenwood, RN, PhD
Clinical Performance Improvement Consultant
Office of Patient Experience, Quality and Clinical Effectiveness
Sutter Health
2200 River Plaza Drive
Sacramento, CA, 95833
United States
Phone: 1 916 712 4597
Fax: 1 916 503 7596
Email: greenwd@sutterhealth.org

Abstract

Background: Type 2 diabetes mellitus is a worldwide challenge. Practice guidelines promote structured self-monitoring of blood glucose (SMBG) for informing health care providers about glycemic control and providing patient feedback to increase knowledge, self-efficacy, and behavior change. Paired glucose testing—pairs of glucose results obtained before and after a meal or physical activity—is a method of structured SMBG. However, frequent access to glucose data to interpret values and recommend actions is challenging. A complete feedback loop—data collection and interpretation combined with feedback to modify treatment—has been associated with improved outcomes, yet there remains limited integration of SMBG feedback in diabetes management. Incorporating telehealth remote monitoring and asynchronous electronic health record (EHR) feedback from certified diabetes educators (CDEs)—specialists in glucose pattern management—employ the complete feedback loop to improve outcomes.

Objective: The purpose of this study was to evaluate a telehealth remote monitoring intervention using paired glucose testing and asynchronous data analysis in adults with type 2 diabetes. The primary aim was change in glycated hemoglobin (A_{1c})—a measure of overall glucose management—between groups after 6 months. The secondary aims were change in self-reported Summary of Diabetes Self-Care Activities (SDSCA), Diabetes Empowerment Scale, and Diabetes Knowledge Test.

Methods: A 2-group randomized clinical trial was conducted comparing usual care to telehealth remote monitoring with paired glucose testing and asynchronous virtual visits. Participants were aged 30-70 years, not using insulin with A_{1c} levels between 7.5% and 10.9% (58-96 mmol/mol). The telehealth remote monitoring tablet computer transmitted glucose data and facilitated a complete feedback loop to educate participants, analyze actionable glucose data, and provide feedback. Data from paired glucose testing were analyzed asynchronously using computer-assisted pattern analysis and were shared with patients via the EHR weekly. CDEs called participants monthly to discuss paired glucose testing trends and treatment changes. Separate mixed-effects models were used to analyze data.

XSL•FO
RenderX

Results: Participants (N=90) were primarily white (64%, 56/87), mean age 58 (SD 11) years, mean body mass index 34.1 (SD 6.7) kg/m2, with diabetes for mean 8.2 (SD 5.4) years, and a mean A_{1c} of 8.3% (SD 1.1; 67 mmol/mol). Both groups lowered A_{1c} with an estimated average decrease of 0.70 percentage points in usual care group and 1.11 percentage points in the treatment group with a significant difference of 0.41 percentage points at 6 months (SE 0.08, t_{159}=–2.87, *P*=.005). Change in medication (SE 0.21, t_{157}=–3.37, *P*=.009) was significantly associated with lower A_{1c} level. The treatment group significantly improved on the SDSCA subscales carbohydrate spacing (*P*=.04), monitoring glucose (*P*=.001), and foot care (*P*=.02).

Conclusions: An eHealth model incorporating a complete feedback loop with telehealth remote monitoring and paired glucose testing with asynchronous data analysis significantly improved A_{1c} levels compared to usual care.

Trial Registration: Clinicaltrials.gov NCT01715649; https://www.clinicaltrials.gov/ct2/show/NCT01715649 (Archived by WebCite at http://www.webcitation.org/6ZinLl8D0).

(J Med Internet Res 2015;17(7):e178) doi:10.2196/jmir.4112

KEYWORDS

telehealth; remote consultation; electronic health records; health records, personal; diabetes mellitus, type 2; self-care; monitoring, physiologic; blood glucose self-monitoring; hemoglobin A1c, glycosylated; eHealth; patient participation

Introduction

In the United States, 9.3% of Americans have diabetes mellitus; of those, 90% to 95% are diagnosed with type 2 diabetes [1]. When uncontrolled, diabetes is the seventh leading cause of death and the leading cause of kidney failure, blindness, and nontraumatic amputations in the United States [1]. Achieving national diabetes outcome targets for blood glucose, blood pressure, and blood fats can decrease complications and improve quality of life [2]. However, research indicates people with diabetes remain at suboptimal glucose control for 2.9 years from patient and provider clinical inertia limiting treatment intensification [3,4]. Self-management of diabetes is a critical component of diabetes care [2] and self-monitoring of blood glucose (SMBG) is an essential self-management behavior [2,5]. Evaluation of SMBG data by primary care providers encourages more frequent medication changes and several studies indicate improved glycemic control [6-8]. Practice guidelines promote the use of SMBG for informing health care providers about glycemic control and providing patient feedback to increase knowledge, self-efficacy, and behavior change [9-11]. Effective SMBG includes structured behaviors such as (1) frequency of glucose testing, (2) participant use and response to glucose data, (3) health care provider data interpretation, and (4) therapy modifications [12,13]. However, there is controversy regarding the benefit of SMBG to improve glycated hemoglobin (A_{1c})—a measure of overall blood glucose control—in persons with noninsulin-treated type 2 diabetes with some systematic reviews reporting no reduction in A_{1c} [14,15]. However, current research incorporating structured monitoring profiles—defining the frequency, intensity, and timing of SMBG—shows significant improvement in A_{1c} [6,7,16,17]. Although there is limited research to document the most effective SMBG profile, paired glucose testing (eg, pairs of glucose results obtained before and after a meal, physical activity, or other event) is one suggested profile [10,18]. However, it is challenging to access glucose data frequently to interpret values and recommend patient actions.

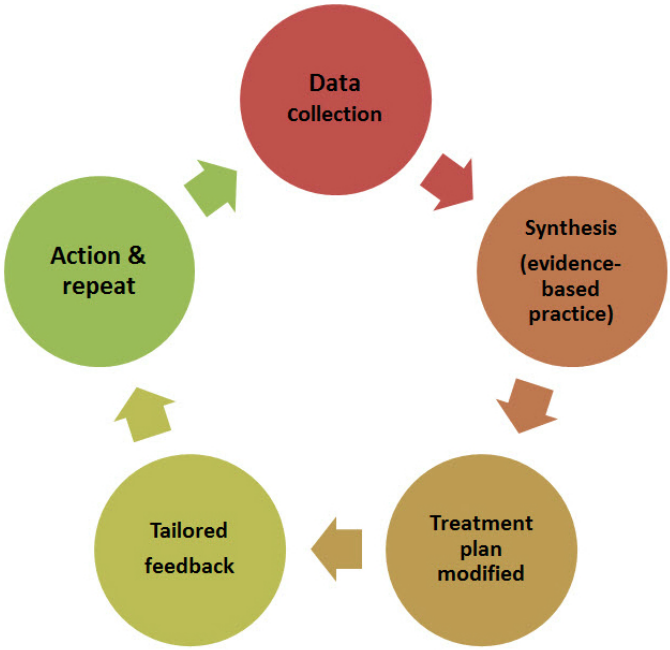
A complete feedback loop—data collection and interpretation combined with feedback to the patient to modify treatment plan—has been associated with improved outcomes [19]. Although the complete feedback loop is an essential component of both SMBG [20] and remote monitoring, there is limited and inconsistent incorporation of SMBG feedback in diabetes management [14,19-21]. Although research has evaluated telehealth remote monitoring glucose data, few studies have incorporated SMBG profiles that provide timely feedback to patients and allow for real-time decision making [22].

In primary care, health care providers are often not prepared to interpret SMBG data, respond to patterns, and implement a complete feedback loop with tailored feedback for behavior change or treatment modifications [14,23]. Diabetes management programs with nurse care coordination [14,23] often include diabetes education provided by certified diabetes educators (CDEs), who are uniquely qualified to analyze SMBG data and problem solve with patients. Incorporating telehealth remote monitoring with CDE support employs the complete feedback loop to improve outcomes. Figure 1 shows the complete feedback loop elements [19]. The patient generates glucose data following targeted education on the elements of structured SMBG. Next, data are analyzed and synthesized by both the CDE and the patient using pattern management and evidence-based guidelines. In collaboration, the CDE and patient agree on modification of the existing treatment plan through active communication and tailored feedback from the CDE. Finally, a new action plan is developed using shared decision making and implemented by the patient and the cycle continues.

Telehealth remote monitoring may improve clinical outcomes, care coordination, engagement, and satisfaction [24,25]. Novel clinical interventions are needed that expand existing paradigms of diabetes care by utilizing telehealth remote monitoring and actionable patient-generated data for timely behavior and treatment changes. The purpose of this study was to test the effectiveness of a telehealth remote monitoring intervention with paired glucose testing for adults with noninsulin-treated type 2 diabetes. The hypothesis was that the intervention would result in a greater change in A_{1c} and improved self-management, self-efficacy, and diabetes knowledge compared to usual care over 6 months.

XSL•FO
RenderX

Figure 1. Complete feedback loop for improved outcomes in diabetes management.



Methods

Study Design

The study was a 2-group randomized clinical trial with 1:1 randomization to usual care or telehealth remote monitoring with paired glucose testing (treatment group). Sample size was determined based on the main outcome: mean change in A_{1c} between treatment and usual care over 6 months. The comparison of usual care (n=39) to treatment participants (n=39) had 80% power to detect a 0.9% difference in A_{1c} between treatment and usual care after 6 months (α=.05, 2-tailed). A 15% additional margin for participant dropout resulted in a sample size of n=45 per group.

Setting

The study was conducted between January and October 2013 in a large health care system in California with an established diabetes management program with telephonic nurse care coordination for diabetes population health management. CDEs proactively telephoned patients with A_{1c} ≥10% (86 mmol/mol) to develop care plans, whereas patients in lower risk groups (A_{1c} 7.5%-9% or 58-75 mmol/mol) called the program if desired. CDEs were trained in motivational interviewing to support behavior change, structured paired glucose testing, pattern management, and medication management. Approximately 7000 patients were enrolled in the diabetes management program at that time and nearly 1500 patients met the inclusion criteria.

Recruitment and Enrollment

Participants were recruited through query of the electronic health record (EHR) and diabetes management database using the *International Classification of Diseases, Ninth Revision* (ICD-9) code 250.02. The inclusion criteria were:

1. Type 2 diabetes diagnosis treated with oral antihyperglycemic medications, noninsulin injectable medications, or lifestyle alone;
2. Participant in the diabetes management program for previous 12 months;
3. Aged between 30 and 70 years;
4. A_{1c} between 7.5% and 10.9% (58-96 mmol/mol) in previous 6 months;
5. Internet or 3G connection with email access;
6. Landline or cellular phone;
7. English speaking; and
8. Primary care provider in health system.

Exclusion criteria identified by medical chart review included:

1. Insulin prescription;
2. Unable to independently self-manage (diagnosis of dementia, severe depression, schizophrenia, or cognitive impairment for previous 12 months); and/or
3. Diagnoses of debilitating stroke, heart failure, end-stage renal disease, or legally blind.

Eligible participants were contacted through mail, email, and telephone (see Figure 2). Consent forms were mailed and emailed to participants. The research team obtained informed consent by telephone and then participants signed consents and returned by postal mail. We estimated a 15% enrollment rate, but 6% was attained. Major reasons for ineligibility were non-English speakers (21%, 8/37), insulin (21%, 8/37), primary care provider not in health system (16%, 6/37), and no Internet access (14%, 5/37).

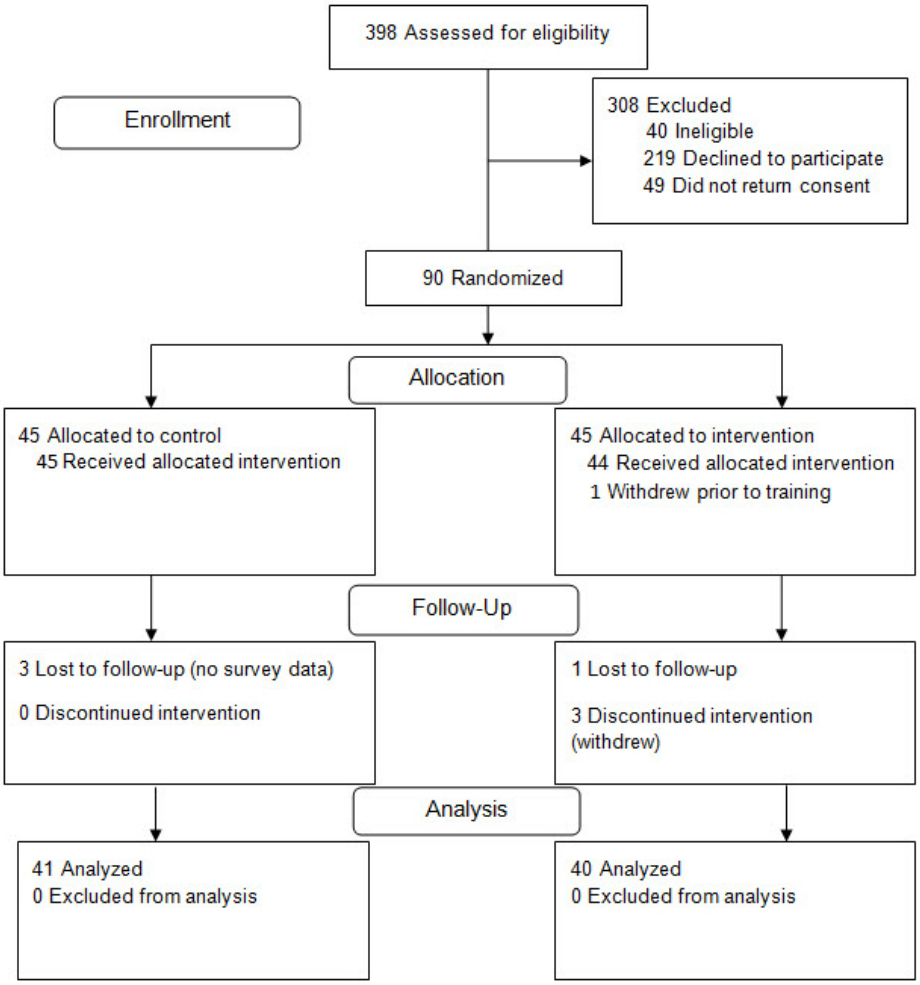
A permuted block, with blocks of 4 and 6, and a computer-generated random number table were utilized for randomization. After participants signed the consent form, the research coordinator assigned sequential study identification (ID) numbers. The investigator matched the ID numbers to the random number table to assign study group. Participants were

XSL•FO
RenderX

notified of group assignment by email after completing online baseline self-assessment questionnaires. Participants randomized to the control group (usual care) were informed to continue in the diabetes management program for usual care. Blinding of participants, providers, and the research team was not possible.

Participants in both groups received a US $10 gift card after completing online questionnaires. A_{1c} tests were ordered every 3 months, as is standard of care when A_{1c} is elevated, then billed to insurance. A_{1c} tests were collected at health system laboratories using similar equipment following standardized procedures. Questionnaires were completed online using the Research Electronic Data Capture (REDCap) database. The study was approved by Sutter Health Central Institutional Review Committee and a Data and Safety Monitoring Board (DSMB) reviewed the study procedures and adverse events.

Figure 2. CONSORT flowchart of enrollment and participant status.



Measures

Primary Outcome

The primary outcome was the difference in mean change in A_{1c} from baseline to 6 months between groups. A_{1c} at recruitment (prestudy A_{1c}), baseline, and 3 and 6 months postprogram was obtained from EHR review. Baseline data were the most recent A_{1c} recorded before study enrollment in the previous 6 months. At the conclusion of the study, the A_{1c} was scheduled within a 3-month time period, approximately 4 weeks before and 8 weeks after the 6-month due date.

Secondary Outcomes

Diabetes Knowledge

Diabetes knowledge was measured using the Diabetes Knowledge Test (DKT) [26], a valid and reliable measure for estimating general understanding of diabetes, including healthy eating and glucose monitoring consisting of 23 multiple-choice items. The first 14 items, appropriate for people not using insulin, were administered to study participants. Scores are measured as the number of correct answers divided by the possible total of 14.

Self-Management

Self-management was measured by the Summary of Diabetes Self-Care Activities (SDSCA) [27], a 12-item self-report questionnaire with subscales: general diet, specific diet, carbohydrate spacing, exercise, monitoring blood glucose, and foot care. For example, participants were asked: “In how many of the past 7 days (0-7) did you check your blood glucose?” Higher scores indicate better self-care behavior.

XSL•FO
RenderX

Self-Efficacy

Self-efficacy was measured by the Diabetes Empowerment Scale short form (DES-SF), an 8-item measure of psychosocial self-efficacy in people with diabetes [28]. Scores ranged from 1-5, with 5 indicating "strongly agree." The mean score of 8 items is reported.

Usual Care/Control Group

Participants in usual care received diabetes education booklets and referral for formal diabetes education as needed. This group continued to receive nurse care coordination including reminders for A_{1c} and health maintenance exams sent by postal mail. The CDEs evaluated SMBG data when reported by participants (no specific monitoring profile was required) and discussed behavior changes with participants by telephone and/or secure messaging and discussed possible medication changes with their primary care provider through the EHR staff messaging tool. CDE contact with the usual care group was documented in the study database.

Treatment Group: Structured Glucose Monitoring

The intervention incorporated a complete feedback loop and all essential elements of structured monitoring. Before the intervention, 6 CDEs attended in-person training sessions on intervention procedures, paired testing, and the goal of implementing a complete feedback loop. Participants in the treatment group attended a 1-hour, small group, in-person training session led by the CDE that included (1) use of the glucose meter, (2) implementation of the complete feedback loop, (3) use of paired glucose testing (frequency and intensity of monitoring), (4) American Diabetes Association (ADA) goals for pre- and postmeal, (5) how to use SMBG data to modify behavior or treatment, (6) expected feedback from CDEs with communication by secure message or phone, and (7) the use of shared decision making to implement the treatment plan [22]. Participants created a "personal experiment" and agreed to check glucose before and 2 hours after the same meal or physical activity for 1 week, and created a behavior change action plan to evaluate changes in SMBG data.

During training, participants were educated on how to use the Care Innovations Guide, a telehealth remote monitoring system approved by US Food and Drug Administration (FDA), which includes an in-home tablet computer connected by Internet or 3G network to the Care Innovations Health Suite online portal (Intel-GE Care Innovations, Roseville, CA, USA). The Care Innovations Guide is connected to the glucometer via USB cables and has a touchscreen for participants to answer daily health session questions. Data are downloaded to the Health Suite for CDEs to access via the Internet. Participants received a OneTouch Ultra 2 glucometer (approved by the FDA), test supplies, and USB cables to keep (Johnson and Johnson-Lifescan Inc, Milpitas, CA, USA) at no charge. Participants returned the Care Innovations Guide when the intervention concluded.

The 84 sequential daily health sessions, designed by the research team, were delivered electronically through the Care Innovations Guide as a text document in the style of a PowerPoint slide or via short video clips. The daily health session started with an audible prompt from the Care Innovations Guide at a time convenient for the participant, then participants completed a glucose check while the glucometer was connected to the Care Innovations Guide via the USB cable. Glucose data were automatically transferred to the Care Innovations Guide at that time. Participants read brief educational content focusing on 1 or 2 key points from the American Association of Diabetes Educators AADE7 [5] self-care behaviors (healthy eating, being active, monitoring, taking medication, problem solving, reducing risks, and healthy coping) a framework for organizing education and structuring behavior change goals. An automated health session reminded participants to evaluate glucose data, using pattern management, every week and to revise or continue their personal experiment for the following week. Participants were assigned a CDE to contact by secure message or phone for diabetes-related questions, instructed to contact their primary care provider for additional health care needs, and to call 911 for emergencies.

Data Review and Nurse Care Coordination

The CDEs reviewed health session and SMBG data in the Health Suite Web portal, stratified by a stoplight system with red indicating missing data or data above or below predetermined thresholds, yellow indicating pending data, and green indicating all data within range. CDEs telephoned participants, at predetermined times, when SMBG data indicated an urgent situation, such as severe hypoglycemia (1 value <50 mg/dL) or hyperglycemia (1 value >450 mg/dL). CDEs also telephoned participants if they reported a change in their feet or a new problem with medication by answering "yes" to health session questions. The Web portal data were reviewed by CDEs during normal business hours Monday through Friday. Data entered during nonbusiness hours were reviewed the following business day. Glucose data were analyzed weekly via software specifically designed for the intervention and evaluated against ADA goals of 80-130 mg/dL before meals, ≤180 mg/dL 2 hours postmeal, and a 30-50 point change between premeal to postmeal. After SMBG analysis, CDEs generated a virtual visit via asynchronous secure messaging through the EHR using the secure message feature. CDEs created a virtual encounter in the EHR, then "copy and pasted" a summary of SMBG pattern analysis data along with personalized feedback and individualized care coordination to reinforce action plans to create the virtual visit for both participants and providers to read (Figure 3). CDEs telephoned participants at weeks 4, 8, and 12 for a 30-minute discussion of SMBG trends, patterns, and goal achievement using motivational interviewing to identify opportunities to improve glucose values. If SMBG data did not improve after 4 weeks, CDEs discussed medication options with patients and/or primary care providers using shared decision making [29]. CDEs incorporated virtual visit data, both preprandial and postprandial glucose, to suggest medication changes, including insulin therapy. Medication changes were ordered by primary care providers via the EHR. Participants were instructed to use paired glucose testing or a monthly, 3-day 7-point glucose profile until the 6-month A_{1c}. CDEs documented patient contact in the study database.

XSL•FO
RenderX

Figure 3. Sample weekly paired glucose testing data analysis, by software designed for the study, and sample message text for feedback to participants through asynchronous secure messaging via the electronic health record.

Date	Premeal glucose, mg/dL (mmol/L)	Postmeal glucose, mg/dL (mmol/L)	Change mg/dL
6/2/2013	203 (11.3)	218 (12.1)	15
6/3/2013	192 (10.7)	227 (12.6)	35
6/4/2013	256 (14.2)	232 (12.9)	-24
6/5/2013	217 (12.1)	274 (15.2)	57
6/6/2013	193 (10.7)	277 (15.4)	84

Message Text:

- Hi, it looks like you have 5 sets of paired tests.
- The pre to post change is below the goal of 50 points 3/5 times; 2 are above goal and average 71.
- 0/5 premeal values are within the goal of 80-130; 5 are above goal and average 212.
- 0/5 postmeal values are within the goal of less than 180; 5 are above goal and average 246.

Statistical Analysis

Mixed-effects models were used to compare mean change over time in primary and secondary outcomes between groups. A_{1c} was measured at baseline and at approximately 3 and 6 months. Change in A_{1c} was evaluated by fitting different growth models. Time was represented in the models by 90-day increments with the estimated change in A_{1c} equaling the amount of change of approximately 3 months. An indicator of group membership was added to the best-fitting growth model to test for differences between groups in A_{1c} at 3 and 6 months and the change in A_{1c} statistically adjusting for prestudy A_{1c}. An indicator variable was added to the model that denoted whether a participant changed medication during the study and tested the effect of medication change on A_{1c} at 3 and 6 months, and on the change in A_{1c}. Finally, a model was fit to test the effect of the number of paired glucose tests on A_{1c} at 3 and 6 months and change in A_{1c} over time controlling for change in medication and prestudy A_{1c}. The effect of the number of paired glucose tests on A_{1c}, accounting for effects due to changes in medication, was studied. Tests used a significance level of $P<.05$ or a 95% confidence interval that excluded zero. SAS version 9.4 (SAS Institute, Inc, Cary, NC, USA) was used to obtain restricted maximum-likelihood estimates using PROC MIXED.

Results

Overview

Table 1 shows baseline characteristics of participants (N=90). The majority were white (64%, 56/87), had diabetes for a mean 8.2 (SD 5.4) years, mean age of 58 (SD 11) years, and mean A_{1c} of 8.3% (SD 1.1; 67 mmol/mol). Participants were highly educated (63%, 55/87 college/postcollege), employed (53%, 46/87), with previous diabetes education (86%, 75/87), hypertension (59%, 51/87), and hyperlipidemia (69%, 60/87). There were no differences between groups at baseline except for self-reported hyperlipidemia (P=.006) in the treatment group. All data were included in an intent-to-treat analysis. There were no serious hyper- or hypoglycemic events or hospitalizations. One participant visited the emergency department unrelated to the study and DSMB determined there were no serious adverse events related to the study.

XSL•FO
RenderX

Table 1. Demographic and key baseline characteristics by group.

Characteristic	Usual care n=45	Treatment n=45
Female, n (%)	19 (21)	23 (25)
Age (years), mean (SD)	57.5 (10.6)	53.9 (10.4)
Years with diabetes (years),[a] mean (SD)	8.1 (5.3)	8.3 (5.5)
Ethnicity[a], n (%)		
Hispanic	8 (9)	6 (7)
White	27 (31)	29 (33)
Black/African American	2 (2)	1 (1)
American Indian	1 (1)	2 (2)
Asian/Pacific Islander	4 (5)	3 (3)
Other	1 (1)	2 (2)
Not reported	0 (0)	1 (1)
Education[a], n (%)		
College	18 (21)	17 (20)
High school	10 (12)	16 (18)
Other	3 (3)	3 (3)
Post college	12 (14)	8 (9)
Employment status[a], n (%)		
Employed	21 (24)	25 (29)
Not employed	6 (7)	6 (7)
Retired	16 (18)	13 (15)
Marital status[a], n (%)		
Married	29 (33)	36 (41)
Single/divorced/widowed	14 (16)	8 (9)
Previous diabetes education[a], n (%)	40 (46)	35 (40)
Computer use[a], n (%)	42 (48)	44 (51)
Type of Internet use[a], n (%)		
Email, yesterday	34 (39)	41 (47)
News, yesterday	24 (28)	34 (39)
Medical, yesterday	9 (10)	14 (16)
Video, yesterday	11 (13)	14 (16)
Treatment[a], n (%)		
Lifestyle	2 (2)	3 (3)
Pills	10 (12)	7 (8)
Pills and lifestyle	27 (31)	30 (35)
Noninsulin injectable	2 (2)	3 (3)
Noninsulin injectable and pills	2 (2)	1 (1)
Comorbidities[a], n (%)		
Heart attack	6 (7)	5 (6)
Coronary heart disease	5 (6)	4 (5)

XSL•FO
RenderX

Characteristic	Usual care n=45	Treatment n=45
Atherosclerosis	3 (3)	2 (2)
Stroke	3 (3)	0 (0)
Hypertension	22 (25)	29 (33)
High cholesterol	24 (28)	36 (41)[b]
Satisfaction with care[a], n (%)		
Strongly agree	11 (13)	6 (7)
Somewhat agree	14 (16)	13 (15)
Neutral	9 (10)	12 (14)
Somewhat disagree	8 (9)	6 (7)
Strongly disagree	1 (1)	7 (8)
Body mass index (kg/m^2), mean (SD)	34.1 (6.6)	34.1 (6.8)
Blood pressure (mm Hg), mean (SD)		
Systolic	128.8 (13.9)	126.9 (13.2)
Diastolic	76.6 (11.0)	77.3 (9.1)
Triglycerides (mg/dL), mean (SD)	175.5 (111.3)	170.5 (112.3)
High-density lipoprotein (mg/dL), mean (SD)	39.8 (10.6)	37.9 (12.2)
Low-density lipoprotein (mg/dL), mean (SD)	92.1 (29.4)	92.8 (28.8)
Cholesterol (mg/dL), mean (SD)	164.4 (35.6)	161 (38)
A_{1c}, mean (SD)		
%	8.2 (1.1)	8.5 (1.1)
mmol/mol	7 (13)	69 (12)

[a] Missing baseline questionnaire data for 3 participants (n=87).

[b] P=.006; 2-tailed P value corresponding to a test of a difference between usual care and treatment groups.

Primary Outcomes

The best-fitting model to describe A_{1c} over time was a quadratic growth model suggesting that the rate of change in A_{1c} was not constant over time. The quadratic model included 3 coefficients. The first two, the intercept and the linear change rate, described A_{1c} level and change in A_{1c} at specific time points. We evaluated these 2 coefficients at baseline and at 3- and 6-months postbaseline. Time was coded to reflect change at approximately 3-month intervals; thus, the linear change rate was the expected rate of change using a 90-day interval. The third coefficient, the quadratic effect, was the acceleration rate that allowed for a nonconstant rate of change over time. Of the 3 coefficients, only the intercept varied across individuals suggesting individual differences in A_{1c} levels over time, but no significant individual differences in the change rate. After fitting a growth model to data, individual scores may correlate between measurements. Model fit for A_{1c} scores was improved by allowing the residuals between measurements to correlate within persons.

Using a quadratic growth model to describe A_{1c} over time, we tested group differences in A_{1c} levels and the linear rates of change at baseline and at 3 and 6 months controlling for prestudy A_{1c}. Comparisons between usual care group and treatment group suggested no difference in mean A_{1c} or in the linear change rate at baseline. An estimated group mean A_{1c} difference of 0.11 (t_{159}=0.63, P=.53) at 3 months and –0.11 (t_{159}=–0.59, P=.55) at 6 months showed no significant differences between groups. At 3 months, the usual care group decreased A_{1c} at a mean rate of –0.35 units (t_{159}=–4.37, P<.001). Between groups, the difference in the change rate of –0.21 (t_{159}=–1.87, P=.06) was not significant, suggesting no difference in the change rate at 3 months. At 6 months, the change rate in A_{1c} for the usual care group of –0.07 (t_{159}=–1.41, P=.16) was not statistically significant, indicating no further improvement in A_{1c} at 6 months. However, the groups differed significantly in the change rate at 6 months, with the treatment group decreasing 0.23 units faster than the usual care group (t_{159}=–2.87, P=.005) (Table 2). Figure 4 shows A_{1c} trajectories for groups over 6 months. Finally, the estimated acceleration rate for the usual care group was 0.14 (t_{159}=4.07, P<.001), suggesting that the change rate increased with time. The estimated acceleration rate for the treatment group was 0.14 + (–0.013)=0.13, although the difference in this coefficient between groups was not significant (t_{159}=–0.26, P=.80) suggesting no difference in this aspect of change in A_{1c} over time.

XSL•FO
RenderX

Table 2. Estimated mean A_{1c} level and instantaneous linear change in A_{1c} at baseline and 3 and 6 months with group differences and prestudy A_{1c} level as a covariate.[a]

Factor	n	Usual care group	Treatment group	Group difference	t-ratio$_{159}$[b]	P
Mean A_{1c} level, % (mmol/mol)						
Baseline	90	8.16 (66)	8.46 (69)	0.30	1.57	.12
3 months	83	7.68 (60)	7.81 (62)	0.11	0.63	.53
6 months	80	7.46 (58)	7.35 (57)	–0.11	–0.59	.55
Prestudy A_{1c} effect[c]		0.52	—	—	5.23	<.001
Instantaneous linear change in A_{1c}[d]						
Baseline	90	–0.62	–0.80	–0.18	–0.94	.35
3 months	83	–0.35	–0.56	–0.21	–1.87	.06
6 months	80	–0.07	–0.31	–0.23	–2.87	.005
Estimated change from baseline to 6 months		–0.70	–1.11	–0.41		
Acceleration rate[e]		0.14	0.12	–0.01	–0.26	.80

[a] Tabled values are maximum-likelihood estimates.

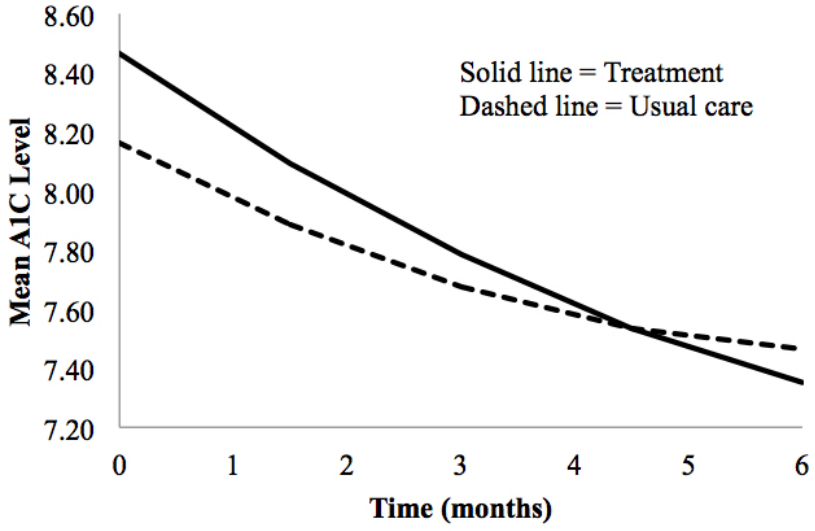
[b] t-ratios are ratios of the estimates to their respective standard errors.

[c] The prestudy A_{1c} effect, a regression coefficient, is the change in A_{1c} when measured during the study for a unit increase in prestudy A_{1c} level.

[d] Instantaneous linear change in A_{1c} reflects the point change in A_{1c} for a 90-day increment.

[e] Acceleration rate is the rate of acceleration of the quadratic growth model.

Figure 4. Estimated A1C trajectories for the usual care and treatment groups from baseline to 6 months.



One usual care participant and 27 participants in the treatment group self-reported medication change, including 4 starting insulin. A change in medication was related to lower A_{1c} at 3 months (t_{157}=–3.42, P<.001 and 6 months (t_{157}=–3.37, P<.001) controlling for treatment effect and prestudy A_{1c}. The effect of a medication change was not significant on the change rate in A_{1c} at 3 or 6 months (see Table 3). Three usual care group participants engaged with CDEs via telephone during the study. Although only 1 participant changed medication, they all decreased their A_{1c} levels.

The paired glucose testing goal was 84 pairs over 12 weeks with actual values ranging from zero to 73 pairs, mean 10.2 (SD 14.4) pairs, and median 21 (IQR 15) pairs. The effect of the number of paired glucose tests was not statistically significant on either the level (t_{156}=–0.99, P=.33) or change rate in A_{1c} at 3 months (t_{156}=–1.82, P=.07) and the level (t_{156}=–1.86, P=.06) or change rate in A_{1c} at 6 months (t_{156}=–1.82, P=.07) (see Table 3).

XSL•FO
RenderX

Table 3. Estimated effects of change in medication and number of paired glucose tests on A_{1c} level and instantaneous linear change in A_{1c} at baseline and at 3 and 6 months.

Predictor	Estimated effect on A_{1c} level[a]			Estimated effect on the instantaneous linear change rate in A_{1c}		
	MLE[b]	*t*-ratio (*df*)[c]	*P*	MLE[b]	*t*-ratio (*df*)[c]	*P*
Medication change						
3 months	–1.05	–3.42 (157)	<.001	0.35	1.81 (157)	.07
6 months	–0.71	–3.37 (157)	<.001	0.35	1.81 (157)	.07
Number of paired tests						
3 months	–0.007	–0.99 (156)	.33	–0.008	–1.82 (156)	.07
6 months	–0.015	–1.86 (156)	.06	–0.008	–1.82 (156)	.07

[a] Models were estimated in a hierarchical fashion.

[b] MLE: maximum-likelihood estimate.

[c] *t*-ratios are ratios of the estimates to their respective standard errors.

Secondary Outcomes

Table 4 shows secondary outcomes by group over 3 months. Both groups showed improvement on average for general diet, specific diet, carbohydrate spacing, and foot care self-management behaviors measured by the SDSCA. The treatment group showed greater improvement in the self-management behaviors of carbohydrate spacing, monitoring glucose, and foot care. Neither group showed improvement in DKT, DES, or SDSCA subscales of exercise, smoking, and medication.

XSL•FO
RenderX

Table 4. Patient-reported secondary outcomes by group.

Secondary outcome[a]	Usual care		Treatment		*t*-ratio (*df*)	*P*
	n	Mean (95% CI)	n	Mean (95% CI)		
DES						
Baseline	42	3.5 (3.3, 3.8)	45	3.8 (3.2, 4.4)	1.32 (85)	.19
3 months	41	3.8 (3.2, 3.3)	40	4.1 (2.8, 5.3)[b]	0.39 (78)	.70
DKT						
Baseline	42	12.0 (11.3, 12.6)	45	12.4 (10.9, 13.9)	0.98 (85)	.33
3 months	41	11.4 (10.1, 12.6)	40	12.1 (9.1, 14.0)[b]	0.61 (78)	.55
SDSCA						
General diet						
Baseline	42	3.7 (3.2, 4.3)	45	3.7 (2.4, 5.0)	–0.07 (85)	.95
3 months	41	4.9 (3.7, 6.1)	40	4.7 (2.0, 7.0)[b]	–0.20 (78)	.84
Specific diet						
Baseline	42	2.9 (2.4, 3.4)	45	3.5 (2.3, 4.7)	1.69 (85)	.09
3 months	41	4.2 (3.0, 5.3)	40	4.6 (1.7, 7.0)[b]	–0.41 (78)	.68
Carbohydrate spacing						
Baseline	42	3.0 (2.3, 3.8)	45	2.7 (1.0, 4.6)	–0.46 (85)	.65
3 months	41	3.9 (2.4, 5.4)	40	4.7 (3.4, 7.0)[b]	2.08 (78)	.04
Exercise						
Baseline	42	2.4 (1.7, 3.1)	45	2.7 (1.1, 4.3)	0.56 (85)	.58
3 months	41	2.6 (1.3, 3.9)	40	3.7 (0.60, 6.8)	1.82 (78)	.07
Medication						
Baseline	42	6.5 (6.0, 7.0)	45	6.2 (4.9, 7.0)[b]	–0.93 (85)	.35
3 months	41	6.4 (5.4, 7.3)	40	6.3 (5.2, 7.0)[b]	0.95 (78)	.34
Monitoring glucose						
Baseline	42	3.6 (2.8, 4.4)	45	3.0 (1.1, 4.8)	–1.17 (85)	.25
3 months	41	3.7 (2.1, 5.3)	40	5.1 (2.6, 7.0)[b]	3.31 (78)	.001
Foot care						
Baseline	42	2.5 (1.8, 3.2)	45	2.5 (0.9, 4.2)	0.04 (85)	.97
3 months	41	3.8 (2.5, 5.2)	40	5.0 (1.7, 7.0)[b]	2.42 (78)	.02

[a] DES: Diabetes Empowerment Scale; DKT: Diabetes Knowledge Test; SDSCA: Summary of Diabetes Self-Care Activities.

[b] Confidence intervals assume symmetry about the mean, reported with the maximum scale score as the upper bound contained within the interval estimate; DES max score=5.0; DKT max score=14.0; SDSCA max score=7.0.

Discussion

Principal Findings

To our knowledge, this is the first telehealth remote monitoring study for type 2 diabetes, within a diabetes management program, that incorporated all essential elements of structured monitoring, including (1) identifying frequency of glucose testing, (2) participant use and response to data, (3) health care provider data interpretation, and (4) therapy modifications [12,13] with paired glucose testing [22]. Structured monitoring created actionable patient-generated data in the context of a complete feedback loop facilitating change in both self-management behavior and treatment. In this study, both groups had improved A_{1c} levels at 3 months without a significant difference between groups in the rate of change (*P*=.06). However, at 6 months, the treatment group continued to have a statistically significant decrease in A_{1c} levels (demonstrating sustained benefit from the intervention), whereas the usual care group participants were no longer improving (*P*=.005). Both groups had lower A_{1c} levels with an estimated average decrease of 0.70 percentage points in the usual care group and 1.11

XSL•FO
RenderX

percentage points in the treatment group, with a significant group difference of 0.41 percentage points at 6 months. Although baseline A_{1c} levels were higher in the treatment group, it was not significantly different from the usual care group. These outcomes are similar to a recent systematic review and meta-analysis that demonstrated a statistically significant and clinically relevant mean difference in A_{1c} of -0.44 percentage points (-4.8 mmol/mol) between treatment and usual care when telehealth was added to usual care in diabetes management [30]. Previously, a difference of -0.50 percentage points in A_{1c} levels between treatment and usual care groups has been reported as clinically meaningful in the literature [14]. Implementation of all complete feedback loop elements (telehealth remote monitoring, structured SMBG, nurse care coordination, and treatment change) is necessary to improve outcomes and future clinical translational research needs to be conducted in the context of the complete feedback loop [20].

Reducing clinical inertia in management of type 2 diabetes was a goal of this intervention [3]. In this study, treatment participants had more self-reported medication changes compared to usual care participants and this was significantly associated with A_{1c} level at both 3 (P<.001) and 6 months (P<.001). The weekly asynchronous virtual visits provided analyzed glucose data to reinforce ADA goals and facilitate pattern management. The paired glucose testing analysis reduced clinical inertia for both patients and providers. Evaluating multiple weeks of SMBG data, often continuously above goal, compared to the usual practice of assessing a single A_{1c} result at 3-month intervals, encouraged medication change. This study is similar to others that incorporated nurse care coordination [25,31,32] to suggest medication changes to primary care providers, but different from those that used nurse practitioners [33,34] who were able to adjust medication independently. Although CDEs suggested medication changes, the primary care providers did not always follow the recommendations or ordered a different medication class. Allowing CDEs to adjust and order medications independently might improve outcomes. Although primary care providers had access to paired glucose testing data analysis, several chose to wait for A_{1c} results before initiating medication change. Targeted primary care provider education before the study may have improved outcomes and increased comfort level with adjusting medications using SMBG data [35].

In the telehealth literature, some studies report an association between the frequency of SMBG and the impact on A_{1c} change [33,36], whereas others report no impact [34]. The STeP study showed structured SMBG data resulted in more frequent and effective treatment changes by primary care providers [6]. The St Carlos [7], ROSSO [37], and PRISMA [17] studies also showed improvement in A_{1c} when structured SMBG data were used to adjust treatment. In this study, more sets of paired glucose testing were not associated with a faster rate of decline in A_{1c} levels (P=.06). However, the frequency of paired glucose testing varied. Identifying opportunities to encourage consistent paired glucose testing may improve outcomes. Research to examine the minimum number of paired glucose tests required to improve outcomes is important. The study was not powered to conduct subanalyses of individuals with consistent weekly paired glucose testing. Newer glucose monitoring technologies, including continuous glucose sensors that collect and store glucose data with minimal fingerstick requirements, may improve primary care provider and patient access to glucose data and reduce clinical inertia. Finally, long-term use of paired glucose testing over 12 months needs to be evaluated. Although A_{1c} data were collected at 6 months, the active intervention ended at 3 months. It is possible that continuing the telehealth remote monitoring would further improve outcomes. A comparative effectiveness study varying the frequency of paired glucose testing and type of remote monitoring feedback is necessary to identify best practices to lower A_{1c} levels [38].

This telehealth study incorporated asynchronous virtual encounters via the EHR to provide feedback on the weekly analysis of paired glucose testing data. This study was similar to others that focused on the use of EHRs for feedback [33,39,40], allowing participants to engage in self-management at a convenient time. This study was also similar to others that analyzed data using treatment algorithms [32], but offered a unique approach by using paired glucose testing data to educate the participant on glucose pattern management and empower participants in self-management.

The treatment group improved in 3 self-management behaviors. Daily education on AADE7 self-care behaviors required the participant to respond through the Care Innovations Guide [5]. The treatment group showed greater improvement in SDSCA self-management subscores of foot care, carbohydrate spacing, and monitoring glucose, all content areas presented in the Care Innovations Guide. The asynchronous virtual visits concentrated on the impact of paired glucose testing and helped participants evaluate carbohydrate quantity in food choices while identifying effect on glucose. This telehealth remote monitoring study is unique in that the participants were taught to analyze paired glucose testing data in the same manner as CDEs, looking for trends and patterns, and adjusting behavior accordingly. Data from a recent systematic review showed that only 31.1% of mobile apps create an opportunity for people with diabetes to share glucose data with providers and even fewer (17.8%) offer an opportunity to analyze data and graphically display results to help the individual learn from their data, whereas only 8.8% of applications supported patients in developing personalized action plans [38]. This study improved on limitations identified in this systematic review by sharing glucose data with the health care team, providing both automated and tailored feedback along with problem solving and goal setting support. Although foot care was not directly an intervention component, foot care was incorporated into the 84 virtual health sessions participants completed through the Care Innovations Guide. Participants improved scores in foot care due to the daily educational content provided through the Care Innovations Guide, which indicates virtual education is a successful option to improve self-management behaviors.

Limitations

Due to challenges with recruitment and saturation of the participant pool, the sample size was small. Results may be different for larger studies powered to detect a smaller change

XSL•FO
RenderX

in A_{1c} levels. The study took place over 6 months, thus long-term outcomes and sustainability over time is unknown. Although there are baseline data from the SDSCA subscale on glucose monitoring, we do not know if the usual care group engaged in paired glucose testing or another profile. However, most participants self-reported not checking glucose on a regular schedule and randomization would account for this issue. A Hawthorne effect is always possible when enrolling individuals motivated to change their self-management behaviors. A delayed entry or a crossover study would address the problem of usual care participants knowing group assignment. The population for the study was an existing diabetes management program, a higher level of usual care than described in most telehealth remote monitoring studies, so between-group differences may be smaller than if compared to typical usual care. Data were not collected on specific dietary or physical activity changes participants made. An online food diary and accelerometer to automatically capture physical activity would create a richer dataset to analyze. Although a treatment fidelity plan was in place, the same CDEs were responsible for treatment and usual care groups, possibly contaminating the usual care group. This intervention did not use mobile technology and was unable to provide real-time feedback, which may have limited outcomes [32].

This study enrolled insured, English-speaking participants in a diabetes management program in a health system. There may be membership bias because the study was conducted in English, yet people affected by diabetes represent multiple ethnic groups, many of whom are underinsured or uninsured. A majority of the participants were white, highly educated, with a strong history of computer use. The results of this study can only be generalized to a similar population. This study should be repeated in populations of lower socioeconomic status without access to sophisticated diabetes management programs. Varying the length of telehealth remote monitoring interventions, including duration and intensity, may help define specific requirements to improve outcomes [41]. There are significant costs associated with SMBG that impact patients and payers and a cost-effectiveness analysis would have provided important information. This intervention required 2 test strips per day whereas Medicare allows for 1 strip per day. Nurse care coordination is expensive and other technology-enabled models of care that facilitate the complete feedback loop and increase patient engagement at a lower cost are needed. Incorporating social media for patient support may reinforce problem solving and behavior change and be less costly. Weitzman and colleagues [42] identified that 31.7% of study participants posted their A_{1c} values on their profile page during a study in the diabetes online community Tudiabetes.org indicating an interest in participating in online peer support for glucose management.

Conclusions

This eHealth clinical trial implemented essential elements of structured monitoring in tandem with telehealth remote monitoring and asynchronous virtual visits through a health system EHR. The complete feedback loop was utilized to educate participants, obtain and analyze actionable SMBG data, provide feedback, and collaborate as a team in the decision-making process. Participants used paired glucose testing data to change behavior, self-reporting an increase in both carbohydrate spacing and glucose monitoring. CDEs, experienced in pattern management and medication adjustment, suggested treatment changes in a reasonable time frame (typically 4 weeks), breaking a link in the cycle of clinical inertia, showing a change in medication was associated with a lower A_{1c} level [3]. At present, this level of nurse care coordination has limited reimbursement. Further research is needed to support eHealth models of care that incorporate remote nurse care coordination by CDEs [43]. Implementing a complete feedback loop in the primary care setting, supported by telehealth remote monitoring and paired glucose testing, improves A_{1c} and self-management behaviors in adults with type 2 diabetes.

Acknowledgments

This research project has research support from the Investigator Initiated Studies program of LifeScan Corporation, Intel-GE Care Innovations, Sutter Institute for Medical Research, The Betty Irene Moore School of Nursing, The Jonas Center for Nursing Excellence, and the University of California Davis, National Center for Advancing Translational Sciences, National Institutes of Health, through grant number UL1 TR 000002. The funders of this study did not play a role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; or in the preparation of the manuscript. Intel-GE Care Innovations did not have veto power over or have say about changing any manuscript text other than the description of the hardware and software they provided. The authors thank Van Luong, Carla Johnson, Joan Olveda, Tom Leonard, and Kimberly Buss, Sutter Health Disease Management Program, for managing the intervention and providing medical and nursing guidance; Brooke Adamska, Sutter Health Disease Management Program, the study coordinator; Dorothy Wullenwaber, Sutter Institute for Medical Research for recruitment; Andrea Hankins, Sutter Institute for Medical Research for database development and data analysis support and Carol Parise, Sutter Institute for Medical Research, for statistical guidance.

Authors' Contributions

DAG was Principal Investigator for the study. DAG, HMY, TSN, SAB, and CCQ were responsible for the conception, design, interpretation of data, revising manuscript for intellectual content, and final approval of published version. DAG was responsible for acquisition of data; DAG and SAB performed data analysis; DAG, SAB, HMY, and CCQ drafted the paper. DAG is the

XSL•FO
RenderX

guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Conflicts of Interest

DAG received research support from the Investigator Initiated Studies program of LifeScan Corporation and Intel-GE Care Innovations. No other conflicts are declared.

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.2 [44].

[PDF File (Adobe PDF File), 145KB - jmir_v17i7e178_app1.pdf]

Multimedia Appendix 2

Greenwood clinical inertia RCT PowerPoint presentation.

[PPTX File, 1MB - jmir_v17i7e178_app2.pptx]

References

1. Centers for Disease Control and Prevention. National Diabetes Statistics Report: Estimates of Diabetes and Its Burden in the United States, 2014. Atlanta, GA: US Department of Health and Human Services; 2014. URL: http://www.cdc.gov/diabetes/data/statistics/2014StatisticsReport.html [accessed 2015-07-03] [WebCite Cache ID 6a9WnQFBA]
2. American Diabetes Association. Standards of medical care in diabetes-2015 abridged for primary care providers. Clin Diabetes 2015 Apr;33(2):97-111. [doi: 10.2337/diaclin.33.2.97] [Medline: 25897193]
3. Khunti K, Wolden ML, Thorsted BL, Andersen M, Davies MJ. Clinical inertia in people with type 2 diabetes: a retrospective cohort study of more than 80,000 people. Diabetes Care 2013 Nov;36(11):3411-3417 [FREE Full text] [doi: 10.2337/dc13-0331] [Medline: 23877982]
4. Ziemer DC, Miller CD, Rhee MK, Doyle JP, Watkins C, Cook CB, et al. Clinical inertia contributes to poor diabetes control in a primary care setting. Diabetes Educ 2005;31(4):564-571. [doi: 10.1177/0145721705279050] [Medline: 16100332]
5. Burke SD, Sherr D, Lipman RD. Partnering with diabetes educators to improve patient outcomes. Diabetes Metab Syndr Obes 2014;7:45-53 [FREE Full text] [doi: 10.2147/DMSO.S40036] [Medline: 24550679]
6. Polonsky WH, Fisher L, Schikman CH, Hinnen DA, Parkin CG, Jelsovsky Z, et al. A structured self-monitoring of blood glucose approach in type 2 diabetes encourages more frequent, intensive, and effective physician interventions: results from the STeP study. Diabetes Technol Ther 2011 Aug;13(8):797-802. [doi: 10.1089/dia.2011.0073] [Medline: 21568751]
7. Durán A, Martín P, Runkle I, Pérez N, Abad R, Fernández M, et al. Benefits of self-monitoring blood glucose in the management of new-onset Type 2 diabetes mellitus: the St Carlos Study, a prospective randomized clinic-based interventional study with parallel groups. J Diabetes 2010 Sep;2(3):203-211. [doi: 10.1111/j.1753-0407.2010.00081.x] [Medline: 20923485]
8. Bonomo K, De SA, Fiora E, Mularoni E, Massucco P, Poy P, et al. Evaluation of a simple policy for pre- and post-prandial blood glucose self-monitoring in people with type 2 diabetes not on insulin. Diabetes Res Clin Pract 2010 Feb;87(2):246-251. [doi: 10.1016/j.diabres.2009.10.021] [Medline: 19954855]
9. Funnell MM, Brown TL, Childs BP, Haas LB, Hosey GM, Jensen B, et al. National standards for diabetes self-management education. Diabetes Care 2012 Jan;35 Suppl 1:S101-S108 [FREE Full text] [doi: 10.2337/dc12-s101] [Medline: 22187467]
10. International Diabetes Federation. 2009. Self-monitoring of blood glucose in non-insulin-treated type 2 diabetes URL: http://www.idf.org/guidelines/self-monitoring [accessed 2015-07-03] [WebCite Cache ID 6Zklf8xjx]
11. Stetson B, Schlundt D, Peyrot M, Ciechanowski P, Austin MM, Young-Hyman D, et al. Monitoring in diabetes self-management: issues and recommendations for improvement. Popul Health Manag 2011 Aug;14(4):189-197. [doi: 10.1089/pop.2010.0030] [Medline: 21323462]
12. Parkin CG, Hinnen D, Campbell RK, Geil P, Tetrick DL, Polonsky WH. Effective use of paired testing in type 2 diabetes: practical applications in clinical practice. Diabetes Educ 2009;35(6):915-927. [doi: 10.1177/0145721709347601] [Medline: 19789359]
13. Polonsky WH, Fisher L, Schikman CH, Hinnen DA, Parkin CG, Jelsovsky Z, et al. Structured self-monitoring of blood glucose significantly reduces A1C levels in poorly controlled, noninsulin-treated type 2 diabetes: results from the Structured Testing Program study. Diabetes Care 2011 Feb;34(2):262-267 [FREE Full text] [doi: 10.2337/dc10-1732] [Medline: 21270183]
14. Klonoff DC, Blonde L, Cembrowski G, Chacra AR, Charpentier G, Colagiuri S, Coalition for Clinical Research-Self-Monitoring of Blood Glucose Scientific Board. Consensus report: the current role of self-monitoring of blood glucose in non-insulin-treated type 2 diabetes. J Diabetes Sci Technol 2011 Nov;5(6):1529-1548 [FREE Full text] [Medline: 22226276]

XSL•FO
RenderX

15. Malanda UL, Welschen LM, Riphagen I, Dekker JM, Nijpels G, Bot SD. Self-monitoring of blood glucose in patients with type 2 diabetes mellitus who are not using insulin. Cochrane Database Syst Rev 2012;1:CD005060. [doi: 10.1002/14651858.CD005060.pub3] [Medline: 22258959]
16. Franciosi M, Lucisano G, Pellegrini F, Cantarello A, Consoli A, Cucco L, et al. ROSES: role of self-monitoring of blood glucose and intensive education in patients with Type 2 diabetes not receiving insulin. A pilot randomized clinical trial. Diabet Med 2011 Jul;28(7):789-796. [doi: 10.1111/j.1464-5491.2011.03268.x] [Medline: 21342243]
17. Bosi E, Scavini M, Ceriello A, Cucinotta D, Tiengo A, Marino R, et al. Intensive structured self-monitoring of blood glucose and glycemic control in noninsulin-treated type 2 diabetes: the PRISMA randomized trial. Diabetes Care 2013 Oct;36(10):2887-2894 [FREE Full text] [doi: 10.2337/dc13-0092] [Medline: 23735724]
18. Schnell O, Alawi H, Battelino T, Ceriello A, Diem P, Felton A, et al. Addressing schemes of self-monitoring of blood glucose in type 2 diabetes: a European perspective and expert recommendation. Diabetes Technol Ther 2011 Sep;13(9):959-965. [doi: 10.1089/dia.2011.0028] [Medline: 21714682]
19. Jimison H, Gorman P, Woods S, Nygren P, Walker M, Norris S, et al. Barriers and drivers of health information technology use for the elderly, chronically ill, and underserved. Evid Rep Technol Assess (Full Rep) 2008 Nov(175):1-1422. [Medline: 19408968]
20. Ceriello A, Barkai L, Christiansen JS, Czupryniak L, Gomis R, Harno K, et al. Diabetes as a case study of chronic disease management with a personalized approach: the role of a structured feedback loop. Diabetes Res Clin Pract 2012 Oct;98(1):5-10. [doi: 10.1016/j.diabres.2012.07.005] [Medline: 22917639]
21. Boren SA, Puchbauer AM, Williams F. Computerized prompting and feedback of diabetes care: a review of the literature. J Diabetes Sci Technol 2009 Jul;3(4):944-950 [FREE Full text] [Medline: 20144344]
22. Greenwood DA, Young HM, Quinn CC. Telehealth remote monitoring systematic review: structured self-monitoring of blood glucose and impact on A1C. J Diabetes Sci Technol 2014 Feb 21;8(2):378-389. [doi: 10.1177/1932296813519311] [Medline: 24876591]
23. Schnell O, Alawi H, Battelino T, Ceriello A, Diem P, Felton AM, et al. Self-monitoring of blood glucose in type 2 diabetes: recent studies. J Diabetes Sci Technol 2013;7(2):478-488 [FREE Full text] [Medline: 23567007]
24. Darkins A, Ryan P, Kobb R, Foster L, Edmonson E, Wakefield B, et al. Care Coordination/Home Telehealth: the systematic implementation of health informatics, home telehealth, and disease management to support the care of veteran patients with chronic conditions. Telemed J E Health 2008 Dec;14(10):1118-1126. [doi: 10.1089/tmj.2008.0021] [Medline: 19119835]
25. Shea S, Weinstock RS, Starren J, Teresi J, Palmas W, Field L, et al. A randomized trial comparing telemedicine case management with usual care in older, ethnically diverse, medically underserved patients with diabetes mellitus. J Am Med Inform Assoc 2006;13(1):40-51 [FREE Full text] [doi: 10.1197/jamia.M1917] [Medline: 16221935]
26. Fitzgerald JT, Funnell MM, Hess GE, Barr PA, Anderson RM, Hiss RG, et al. The reliability and validity of a brief diabetes knowledge test. Diabetes Care 1998 May;21(5):706-710. [Medline: 9589228]
27. Toobert DJ, Hampson SE, Glasgow RE. The summary of diabetes self-care activities measure: results from 7 studies and a revised scale. Diabetes Care 2000 Jul;23(7):943-950 [FREE Full text] [Medline: 10895844]
28. Anderson RM, Fitzgerald JT, Gruppen LD, Funnell MM, Oh MS. The Diabetes Empowerment Scale-Short Form (DES-SF). Diabetes Care 2003 May;26(5):1641-1642. [Medline: 12716841]
29. Inzucchi SE, Bergenstal RM, Buse JB, Diamant M, Ferrannini E, Nauck M, et al. Management of hyperglycemia in type 2 diabetes, 2015: a patient-centered approach: update to a position statement of the American Diabetes Association and the European Association for the Study of Diabetes. Diabetes Care 2015 Jan;38(1):140-149. [doi: 10.2337/dc14-2441] [Medline: 25538310]
30. Marcolino MS, Maia JX, Alkmim Maria Beatriz Moreira, Boersma E, Ribeiro AL. Telemedicine application in the care of diabetes patients: systematic review and meta-analysis. PLoS One 2013 Nov;8(11):e79246 [FREE Full text] [doi: 10.1371/journal.pone.0079246] [Medline: 24250826]
31. Wakefield BJ, Holman JE, Ray A, Scherubel M, Adams MR, Hillis SL, et al. Effectiveness of home telehealth in comorbid diabetes and hypertension: a randomized, controlled trial. Telemed J E Health 2011 May;17(4):254-261. [doi: 10.1089/tmj.2010.0176] [Medline: 21476945]
32. Quinn CC, Shardell MD, Terrin ML, Barr EA, Ballew SH, Gruber-Baldini AL. Cluster-randomized trial of a mobile phone personalized behavioral intervention for blood glucose control. Diabetes Care 2011 Sep;34(9):1934-1942 [FREE Full text] [doi: 10.2337/dc11-0366] [Medline: 21788632]
33. Tang PC, Overhage JM, Chan AS, Brown NL, Aghighi B, Entwistle MP, et al. Online disease management of diabetes: engaging and motivating patients online with enhanced resources-diabetes (EMPOWER-D), a randomized controlled trial. J Am Med Inform Assoc 2013 May 1;20(3):526-534 [FREE Full text] [doi: 10.1136/amiajnl-2012-001263] [Medline: 23171659]
34. Stone RA, Rao RH, Sevick MA, Cheng C, Hough LJ, Macpherson DS, et al. Active care management supported by home telemonitoring in veterans with type 2 diabetes: the DiaTel randomized controlled trial. Diabetes Care 2010 Mar;33(3):478-484 [FREE Full text] [doi: 10.2337/dc09-1012] [Medline: 20009091]

XSL•FO
RenderX

35. Polonsky WH, Jelsovsky Z, Panzera S, Parkin CG, Wagner RS. Primary care physicians identify and act upon glycemic abnormalities found in structured, episodic blood glucose monitoring data from non-insulin-treated type 2 diabetes. Diabetes Technol Ther 2009 May;11(5):283-291. [doi: 10.1089/dia.2008.0087] [Medline: 19425876]
36. McMahon GT, Gomes HE, Hickson HS, Hu TMJ, Levine BA, Conlin PR. Web-based care management in patients with poorly controlled diabetes. Diabetes Care 2005 Jul;28(7):1624-1629 [FREE Full text] [Medline: 15983311]
37. Kempf K, Tankova T, Martin S. ROSSO-in-praxi-international: long-term effects of self-monitoring of blood glucose on glucometabolic control in patients with type 2 diabetes mellitus not treated with insulin. Diabetes Technol Ther 2013 Jan;15(1):89-96. [doi: 10.1089/dia.2012.0213] [Medline: 23194054]
38. Arnhold M, Quade M, Kirch W. Mobile applications for diabetics: a systematic review and expert-based usability evaluation considering the special requirements of diabetes patients age 50 years or older. J Med Internet Res 2014;16(4):e104 [FREE Full text] [doi: 10.2196/jmir.2968] [Medline: 24718852]
39. Shea S, Weinstock RS, Teresi JA, Palmas W, Starren J, Cimino JJ, et al. A randomized trial comparing telemedicine case management with usual care in older, ethnically diverse, medically underserved patients with diabetes mellitus: 5 year results of the IDEATel study. J Am Med Inform Assoc 2009;16(4):446-456 [FREE Full text] [doi: 10.1197/jamia.M3157] [Medline: 19390093]
40. Carter EL, Nunlee-Bland G, Callender C. A patient-centric, provider-assisted diabetes telehealth self-management intervention for urban minorities. Perspect Health Inf Manag 2011;8:1b [FREE Full text] [Medline: 21307985]
41. Stone RA, Sevick MA, Rao RH, Macpherson DS, Cheng C, Kim S, et al. The Diabetes Telemonitoring Study Extension: an exploratory randomized comparison of alternative interventions to maintain glycemic control after withdrawal of diabetes home telemonitoring. J Am Med Inform Assoc 2012;19(6):973-979 [FREE Full text] [doi: 10.1136/amiajnl-2012-000815] [Medline: 22610495]
42. Weitzman ER, Kelemen S, Quinn M, Eggleston EM, Mandl KD. Participatory surveillance of hypoglycemia and harms in an online social network. JAMA Intern Med 2013 Mar 11;173(5):345-351. [doi: 10.1001/jamainternmed.2013.2512] [Medline: 23400234]
43. Gee PM, Greenwood DA, Paterniti DA, Ward D, Miller LM. The eHealth Enhanced Chronic Care Model: a theory derivation approach. J Med Internet Res 2015 Apr;17(4):e86 [FREE Full text] [doi: 10.2196/jmir.4067] [Medline: 25842005]
44. Eysenbach G, Consort- E. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

AADE: American Association of Diabetes Educators
ADA: American Diabetes Association
CDE: certified diabetes educators
DES: Diabetes Empowerment Scale short form
DKT: Diabetes Knowledge Test
DSMB: Data and Safety Monitoring Board
EHR: electronic health record
SDSCA: Summary of Diabetes Self-Care Activities
SMBG: self-monitoring of blood glucose

Edited by G Eysenbach; submitted 22.12.14; peer-reviewed by K Pal, K Fitzner; comments to author 12.03.15; revised version received 08.04.15; accepted 24.05.15; published 21.07.15.

Please cite as:
Greenwood DA, Blozis SA, Young HM, Nesbitt TS, Quinn CC
Overcoming Clinical Inertia: A Randomized Clinical Trial of a Telehealth Remote Monitoring Intervention Using Paired Glucose Testing in Adults With Type 2 Diabetes
J Med Internet Res 2015;17(7):e178
URL: http://www.jmir.org/2015/7/e178/
doi:10.2196/jmir.4112
PMID:26199142

©Deborah A Greenwood, Shelley A Blozis, Heather M Young, Thomas S Nesbitt, Charlene C Quinn. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 21.07.2015. 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

XSL•FO
RenderX

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

Online Recruitment Methods for Web-Based and Mobile Health Studies: A Review of the Literature

Taylor S Lane[1], BA; Julie Armin[1], PhD; Judith S Gordon[1], PhD

University of Arizona, Department of Family and Community Medicine, Tucson, AZ, United States

Corresponding Author:
Taylor S Lane, BA
University of Arizona
Department of Family and Community Medicine
1450 N Cherry Ave
Tucson, AZ, 85719
United States
Phone: 1 520 626 6452
Fax: 1 520 626 6134
Email: taylorlane@email.arizona.edu

Abstract

Background: Internet and mobile health (mHealth) apps hold promise for expanding the reach of evidence-based health interventions. Research in this area is rapidly expanding. However, these studies may experience problems with recruitment and retention. Web-based and mHealth studies are in need of a wide-reaching and low-cost method of recruitment that will also effectively retain participants for the duration of the study. Online recruitment may be a low-cost and wide-reaching tool in comparison to traditional recruitment methods, although empirical evidence is limited.

Objective: This study aims to review the literature on online recruitment for, and retention in, mHealth studies.

Methods: We conducted a review of the literature of studies examining online recruitment methods as a viable means of obtaining mHealth research participants. The data sources used were PubMed, CINAHL, EbscoHost, PyscINFO, and MEDLINE. Studies reporting at least one method of online recruitment were included. A narrative approach enabled the authors to discuss the variability in recruitment results, as well as in recruitment duration and study design.

Results: From 550 initial publications, 12 studies were included in this review. The studies reported multiple uses and outcomes for online recruitment methods. Web-based recruitment was the only type of recruitment used in 67% (8/12) of the studies. Online recruitment was used for studies with a variety of health domains: smoking cessation (58%; 7/12) and mental health (17%; 2/12) being the most common. Recruitment duration lasted under a year in 67% (8/12) of the studies, with an average of 5 months spent on recruiting. In those studies that spent over a year (33%; 4/12), an average of 17 months was spent on recruiting. A little less than half (42%; 5/12) of the studies found Facebook ads or newsfeed posts to be an effective method of recruitment, a quarter (25%; 3/12) of the studies found Google ads to be the most effective way to reach participants, and one study showed better outcomes with traditional (eg in-person) methods of recruitment. Only one study recorded retention rates in their results, and half (50%; 6/12) of the studies recorded survey completion rates.

Conclusions: Although online methods of recruitment may be promising in experimental research, more empirical evidence is needed to make specific recommendations. Several barriers to using online recruitment were identified, including participant retention. These unique challenges of virtual interventions can affect the generalizability and validity of findings from Web-based and mHealth studies. There is a need for additional research to evaluate the effectiveness of online recruitment methods and participant retention in experimental mHealth studies.

(J Med Internet Res 2015;17(7):e183) doi:10.2196/jmir.4359

KEYWORDS

mHealth; Internet health; online recruitment; apps; social media; review

XSL•FO
RenderX

Introduction

Background

Chronic health conditions have replaced acute diseases as the leading causes of both illness and death in the United States [1]. While overall mortality rates have declined, chronic health conditions are becoming more common [1]. There is a great need for empirically based treatment and support for these conditions [2]. While the accessibility and availability of health information is limited when it is delivered solely in medical settings, the Internet and mobile health (mHealth) apps hold promise for expanding the reach of evidence-based health interventions and information [3]. Both tools have opened up a new audience for health information [4]. According to a recent survey, approximately 284.5 million people in the United States (90% of the population) have at least one mobile phone [5]. An estimated 58% of mobile phone users have a smartphone, about 31% of mobile phone owners use their phones to look for health or medical information, and approximately 35% of mobile phone users download apps to track or manage their health [6,7]. A similar report found that 72% of adults look on the Internet for health information [8]. Sixty-five percent of Internet users say they are better informed about health because of Internet and mobile phone use, and 44% report that these technologies have greatly helped their ability to get this information [9]. mHealth apps hold power in their accessibility and reach [10], as well as their ability to lower the costs associated with relaying information, support, and assistance to those who need it [11]. While in-person interventions and consultations produce the highest efficacy, mobile technologies may address the disparity between supply and demand for health-related services and information, thus creating a potentially lower-cost solution to bridging that gap [4,12]. In addition, mHealth apps and Web-based interventions provide flexibility that in-person interventions are unable to provide, such as around-the-clock access to information and personalized feedback [4,13,14]. The benefits of this technology are promising but are in the development stage. However, mHealth is predicted to continue its growth as the technology becomes more pervasive in the US population [3].

Although research has been conducted for many years on Web-based health interventions, research in the area of mHealth is nascent. There are few published studies of mHealth apps designed using theoretical frameworks [4,14-16]. There are a large number of smoking cessation apps available for Apple, Android, and Windows Phone platforms, but the vast majority do not include basic evidence-based practices and are found to have low levels of adherence, meaning that participants quit using the app over time [17,18]. In addition, a majority of Web-based smoking cessation interventions are not evidence-based [19]. There is a clear need for evaluation of Web-based and mHealth interventions to ensure that public health problems, such as smoking, can be addressed with evidence-based tools [17,19].

Web-based interventions and mHealth apps are fundamentally different, but research in these areas can experience similar challenges. There are special issues involved with conducting virtual interventions. Research on mHealth apps and Web-based interventions is difficult due to unique challenges associated with recruiting, enrolling, and retaining participants [4,13]. Recruitment and retention of participants in these studies directly affects sample size, which determines the power and significance of the study [14,20]. If recruitment goals are not met in the initial period, it can negatively influence the chance of finding an effect or, if an extension on the recruitment period is required, it can expand study costs considerably [20]. Participant retention is another main concern for studies using mHealth interventions [4]. Some claim attrition may be due to early interest in the novelty of mHealth apps, which then fades as the innovation wears off [4], while others suggest the lack of personal contact can lead to higher dropout rates [21]. Attrition is inevitable, especially in lifestyle interventions such as diet or smoking cessation where dropout rates less than 20% are rarely achieved, but excessive attrition can reduce study power, increase bias, and lower generalizability [22].

Challenges in enrolling participants have led researchers to look towards alternative methods for identifying potential participants [23,24]. Web-based strategies for attracting participants, such as the use of Internet advertising, email invites, craigslist, online message boards, and more recently, social media, have been explored by researchers in order to find general and specific populations for studies [24]. Some studies indicate that online strategies can be effective at reaching a larger number of potential participants, as well as improving affiliated costs [4,13,24]. Online recruitment can cast a broader net than traditional recruitment, extending to previously hard-to-reach populations, such as young adults, racial/ethnic minorities, and those with low education attainment [25,26]. In addition, Internet-based recruitment can be affordable and reach a wider, more diverse population, thus increasing generalizability [27,28]. However, there is lack of consistency in reporting these population variables in studies employing these methods [29].

In a recently completed study (National Cancer Institute Grant #R41CA162502; J Gordon, PI), our research team developed and evaluated a theory-based mHealth app to improve medication adherence and provide behavioral support for tobacco cessation. Our team used online and in-app methods of recruitment with varying degrees of success. The research team, currently conducting another study (National Cancer Institute Grant #R21CA174639; J Gordon, PI) to test a multi-behavioral smoking, diet, and physical activity mHealth intervention, sought to identify the most effective methods for recruiting and retaining research participants. Our first-hand experience, as well as the desire to formulate recommendations for other researchers, prompted an analysis of the literature on methods of attracting and retaining participants in mHealth studies.

Variables of Interest

The cost of the recruitment method used and participant retention rates were our primary variables of interest. Recruitment cost can be important with a limited budget in order to find the largest, most representative recruitment sample with the allotted funds. Participant retention rates were a primary focus due to our suspicion that the lesser amount of investment

XSL•FO
RenderX

required from participants in virtual research negatively affects retention [4,21]. In addition, the broad reach achieved by Web-based recruitment might bring in people who are unfamiliar with the research environment and are unaware of the extent of the commitment they are making. This might later affect the number of participants who enroll and complete all interventions. This predicted higher attrition rate was evident in our own experience, as well as in the literature on Web-based and mHealth studies [4,21]. Participant demographics were included in the review to allow for comparison to "traditional" methods of recruitment. The "traditional" recruitment methods are defined as telephone, newspapers, radio, TV, flyers, or word-of-mouth. Other variables included were duration of recruitment and intervention because of their influence on participant retention and engagement, which are important factors in the generalizability and success of mHealth research [4,22].

Methods

Search Strategy

A review of the literature published between 2004 and 2014 was conducted from August 2014 to October 2014, using the search terms "mHealth", AND "social media", AND "health behavior change", AND "online recruitment", OR "recruitment", in the databases PubMed, MEDLINE, Cumulative Index of Nursing and Allied Health Literature (CINAHL), PsycINFO, and EbscoHost. Articles' reference lists were also screened for relevant literature. Two searches of the literature were conducted, with the first using more stringent selection criteria. The goal of the initial search was to focus solely on mHealth research. The second search was conducted with more expansive selection criteria to include Web-based interventions due to minimal results from the first search.

Phase I Selection Criteria

The initial review was limited to articles that were published in English, had an abstract available, and reported on research about behavioral interventions to improve health. While certain demographic components such as age can affect recruitment results [23], there were no specifications for age or race of participants included in the review. The first author (TL) screened the abstracts and titles of relevant articles for eligibility. Only peer-reviewed papers were included. Mobile health articles that detailed their recruitment process in their methods section were selected (TL). These articles reported online recruitment methods and outcomes from current or past experimental studies, retention rates, and costs related to online recruitment. This initial search produced only one publication [29].

Phase II Selection Criteria

The first search highlighted the novelty of the mHealth research field. Therefore, an additional, broader search was conducted to see if any Web-based studies, in addition to mHealth studies, published articles detailing their recruitment processes and outcomes. This search utilized the same keywords and similar study selection criteria. The only change was expanding the criteria to include Web-based studies, as well. This second phase search gave the researchers a better understanding of what types of articles were being published about online recruitment and what types of information were being reported. Upon conclusion of the broader search, 11 peer-reviewed publications were included for the literature review. This created a total of 12 peer-reviewed articles [24-27,29-36].

Descriptive Characteristics of the Studies

The following data were collected from each article and recorded in an Excel spreadsheet: (1) authors of the publication, (2) year of publication, (3) article digital object identifier (DOI) number, (4) health domain, (5) study design, (6) number of participants enrolled in the study, (7) general demographics of participants, (8) method of recruitment, (9) total time duration of recruitment, (10) total cost of advertising, as well as cost per click if applicable, (11) percentage retention of participants, (12) study or intervention duration, and (13) most effective methods of recruiting, including cost and number of participants, found by the study.

Results

Description of Studies

Overview

Four of the articles (33%) were published in 2014, three (25%) were published in 2013, and the rest were published between 2006 and 2012. The most common health category was tobacco use/smoking cessation (58%, 7/12). Other categories included mental health, general health, and human immunodeficiency virus (HIV). One third (4/12) of the studies used a randomized control trial as their study design, and one quarter (3/12) used a comparison study design. Other designs included exploratory design, pilot studies, and feasibility trials. The most common participant age group recruited was 18-25 years old. One study included participants between 16-25 [25], and all other research employed broad criteria, requiring participants to be older than 18 years of age. The studies targeting users under 25 used Facebook advertising only, whereas all other studies used a variety of online and traditional methods. Examples of online recruiting methods consist of paid media, including search engine advertising, and earned media, such as posting on various websites and craigslist. A majority (83%, 10/12) of the studies recorded participant engagement in some way. An example of engagement could be the number of page or icon clicks within an app in an mHealth study (see the Participant Engagement and Retention section for a more in-depth explanation).

Recruitment Method

Five of the articles used only one recruitment tool (42%), which was Facebook. These articles specified that they used Facebook, as they were targeting only a small demographic (eg, individuals 18-25). Three of these articles [30-32] used different types of ads through Facebook (17%), some more successful than others. The first article, which looked at depression, found that ads that were closely aligned with the content of the research and used wording regarding a problem to solve were more successful than the other ads [30]. The second article, which looked at smoking cessation, found that newsfeed ads were more successful. They also found that simple images of cigarettes, the study logo, and general informational messaging were the

XSL•FO
RenderX

most successful compared to more complex images and intricate targeted messaging [31]. The third article, which looked at tobacco use, found that Facebook ads were more affordable than previous methods used. They noted success of ads as being dependent upon Facebook's approval of the ad [32]. A quarter (3/12) of the articles used multiple methods of Web-based recruitment [27,33,34]. One study advertised through search term query results on multiple search engines [34], and the other two studies used a combination of various websites, search engines, and social media outlets for recruitment purposes [27,33]. One third of the studies (4/12) used a combination of traditional methods and Web-based methods of recruitment and compared them side-by-side for effectiveness [24,26,29,35]. Two studies [24,35] used Facebook ads as the Web-based method (17%) and compared it to the use of flyers and newspaper ads. One study [26] used national websites, local websites, and search engines, and compared them to the use of billboards, TV and radio ads, and direct mail. One study [29] used health websites (eg, WebMD.com), Google AdWords, Facebook, and Twitter, and compared them to TV, radio, newspaper, email, and word of mouth.

Recruitment Duration and Participant Numbers

Recruitment duration was deemed to be an important factor due to its direct impact on the costs associated with recruitment methods. The majority of articles (67%, 8/12) reported an average recruitment duration of approximately 5 months (range 7 weeks-7 months). Of the studies that recruited for less than a year, half noted two trial periods of recruitment where advertising was changed in some way after the first trial. Two studies (17%) noted the reason for the split in recruitment period as a way to evaluate the effectiveness of different types of ads [30,31]. Articles that specified recruitment duration of less than 1 year had an average of 468 participants. Four articles (33%) reported an average duration of 17 months (range 13-23 months). Studies that recruited for more than 1 year had an average of 3199 participants.

Participant Eligibility

Eligibility criteria were diverse across articles, except for age (>18 years old), which was common across studies (92%, 11/12). Only one study recruited participants under age 18 [25]. Eligibility criteria were specific to the health domain on which the study focused. A majority of the tobacco cessation studies specified eligibility criteria (86%, 6/7), including English literacy, current resident of the United States, some form of Internet access, current smoking (usually defined by number of cigarettes smoked within a specific time frame), and not having used a specific tobacco cessation website or other intervention. Half (6/12) of all articles specified geographic location as an eligibility requirement, usually the country in which the study was being conducted. Half (6/12) of all articles also had intervention-related criteria as eligibility requirements. Examples of intervention-related criteria include access to the Internet for some period of time in a week or not already receiving treatment for the health problem specified in the study. One article did not include any eligibility requirements [26].

Recruitment Costs

Cost was recorded in several ways: as overall recruiting cost, cost-per-click for paid advertising on social media sites and search engines, cost per participant or per completed survey, and/or direct cost related to the specific type of recruitment tool used (eg, Facebook ads). A majority (80%, 4/5) of the studies that used Facebook ads as their sole method of recruitment reported that the advertising features on Facebook helped their recruitment (Table 1). These four studies used Facebook ads successfully to recruit their ideal number of subjects within their budget and reported the ability to target a specific population or demographic as an important factor [25,30-32]. One study that compared Internet advertisements to craigslist and email invitations found Facebook to be the most effective type of Internet ad as it recruited the most participants in the 18-25 year old age range [33]. However, one study that used Facebook as the sole method of recruiting found this strategy was not able to generate sufficient participation for conducting large sample surveys [36]. Two Web-based recruitment studies (17%) did not specify a target age range and used Google ads to successfully to recruit ideal participant numbers while staying within budget [27,34]. One compared advertising on Google to other search engines [34], and the other compared advertising on Google to social media outlets and online forums [27]. Certain search terms, such as those related to seeking a depression test or information about symptoms of depression [27], and "quit smoking" [34] were more effective in converting number of ad views to enrolled participants in both of these studies.

The studies that compared Web-based methods and traditional methods of recruitment reported various findings. One study recorded the highest yield of participants from Facebook, but found online advertising to cost twice as much per participant than using newspaper advertising, fliers, and word of mouth [24]. Another comparison study achieved the highest yield of participants from print advertisements, and found lower cost per participant from print ads and flyers than from Web-based methods [32]. However, this study noted that social networking sites were more likely to reach typically unrepresented populations [32]. The third comparison study reported that Google produced the highest yield of participants with the least cost, followed by news stories on medical Internet media (eg, websites such as WebMD) and traditional media [29]. This study also noted that using paid advertising and free posting on social media was the least cost effective [29].

XSL•FO
RenderX

Table 1. Summary of studies—Study type, participants, and recruitment.

Authors (year)	Study type	N	Participant eligibility	Method of recruitment	Recruitment duration	Recruitment costs
Batterham (2014)	Systematic Investigation	1893	Australian, 18 years old	Facebook advertising: First period—4 ads with direct link (2x2 factorial design-"problem" vs "positive" terminology and "altruistic" vs "self-gain" terminology). Second period-Facebook page that you "liked" and page showed visible links to the survey site. Facebook compared to previously completed study which used postal and telephone surveys.	First period recruited for 1 month. Second period recruited for 2 months.	Direct link cost $9.82 per completed survey. Survey linked to the Facebook page cost $1.51 per completed survey.
Fenner (2012)	Exploratory study	278	Female, 16-25, live in Victoria, Australia, willing to complete health survey	Facebook advertising	4 months	Cost-per-click: $0.67. Cost-per-participant: $20.14
Frandsen, Walters, & Ferguson (2013)	Multisite randomized trial	266	Older than 18, smoke more than 10 cigarettes a day for the last 3 years at least, not enrolled in another smoking cessation trial in the last 3 months, highly motivated to quit (>75 on 100 pt scale)	Two methods: Traditional—flyers, newspaper ads Online—Facebook ads	18 months	Average cost-per-click was AUD $0.95. Cost-per-participant was AUD $42.34. Newspaper cost-per-participant was AUD $21.52
Graham, Bock, Cobb, Niaura, & Abrams (2006)	Randomized controlled trial	764	Older than 18, smoking 5 or more cigarettes a day, no prior QuitNet.com use	Major search engines (AOL, MSN, Yahoo, Google), using an active user-intercept protocol	At least 7 months (recruitment was ongoing at published time)	[Not reported]
Graham, Milner, Saul, & Pfaff (2008)	Comparison of different media campaigns	9,655	Not reported	Two methods: Traditional—billboards, TV, radio, direct mail, physician detailing Online—national websites (banner ads), local websites (banner ads), paid search ads (per click basis)	One 6-month period, one 3-month period	First phase: $35 per participant (with a 9% conversion rate). Second phase: $38 per participant (with a 7% conversion rate)
Heffner, Wyszynski Comstock, Mercer, & Bricker (2013)	Pilot study	222	Older than 18 years, smoked at least five cigarettes daily for the past year, want to quit in the next 30 days, willing to be randomly assigned to either group, lives in the United States, has weekly access to the Internet, English literate, not participating in other smoking cessation interventions, and never used Smokefree.gov website	Two methods: Traditional—Standard media (news coverage on TV, radio, newspaper ads), emails, word of mouth. News coverage on online—medical Internet media, Google AdWords, Facebook advertising and free posts Twitter posts	10 weeks	Total cost: $9,429.83; Direct cost from Facebook: $1,250; Direct cost from Google: $3,320.53; Direct cost from press releases: $1,250; Cost-per-participant: $42.48.

XSL•FO
RenderX

Authors (year)	Study type	N	Participant eligibility	Method of recruitment	Recruitment duration	Recruitment costs
Morgan, Jorm, & Mackinson (2013)	Randomized controlled trial	1326	18 years, from Australia, New Zealand, UK, Ireland, Canada, or US, access to Internet once a week, not getting help for depression already	Search engine advertising, Facebook ads, Forum posts, posts on relevant websites and online newsletters	14 months	Google keyword costs: Cost-per-click=AUD $0.09, cost-per-participant=AUD $10.75 (click-through rate of 6%); Google display network advertising: Cost-per-click= AUD$0.13, cost-per-person= AUD$14.71; Facebook costs: Cost-per-click=AUD $0.62, cost-per-participant=AUD $19.89 (click-through rate of 0.05%)
Ramo, Hall, & Prochaska (2010)	Comparison of three recruitment methods	201	18-25 years of age, English literate, smoked at least one cigarette in past 30 days	craigslist, Internet ads through Adbrite (2 banner ads and 1 text ad), email invitations	6 months	craigslist: Free to post (estimated $0.66 per participant for time spent); Adbrite: Cost-per-participant=$20.86 (Charge every 1000 impression); SSI (online sampling service)=$19.24 per completed survey
Ramo & Prochaska (2012)	Investigation of Facebook as a recruitment mechanism	1548	18-25 years of age, live in the United States, English literate, smoke at least one cigarette in the past 30 days	Facebook's Ad program	13 months	Cost-per-click: $0.45; cost-per-completed survey: $4.28; Overall cost: $6,628.24
Ramo, Rodriguez, Chavez, Sommer, & Prochaska (2014)	Investigation of recruitment campaigns	79	18-25 years of age, English literate, go on Facebook 4 or more days a week, smoked 100 or more cigarettes in their lives, currently smoke one per day on 3 or more days of the week, access to camera required for bioconfirmation of nonsmoking.	Facebook's Ads Manager program- Newsfeed ads and ads on the right column of the page	5 different standard ads, 2 sponsored stories, and 3 promoted posts were up for 3 weeks, 16 ads with a picture and text combination were posted for 7 weeks	Cost-per-click: $0.34; Overall cost: $2.024; Cost-per-participant: $8.80
Raviotta, Nowalk, Lin, Huang, & Zimmerman (2014)	Human Papillomavirus Vaccine Trial	220	18-25 years of age, male, fewer than five lifetime sexual partners, no history of HPV infection or vaccination, no autoimmune disease nor immunosuppression, no hospitalization in past year, no receipt of blood products or immunoglobulins within 90 days, no participation in other drug studies within 30 days, and no receipt of other vaccines within 8 days.	Two methods: Traditional—flyers, email, student newspaper ads, class announcements Online—Facebook ads	One 3-month period, one 5-month period	Facebook: Direct costs= $4,820, Cost-per-click= $1.24, cost-per-person= $110. Print ads and Flyers: Direct costs= $6,758, Cost-per-participant= $61
Valdez, Guterbock, Thompson, Reilly, Menefee, Bennici, Williams, & Rexrode (2014)	Feasibility trial	87	Study 1: 18 years or older, identify as Filipino, live in the United States; Study 2: 18 years or older, US citizen, diagnosed with type 2 diabetes, Facebook user	Already-established Facebook group	Study 1: 5 months; Study 2: 2 months	Study 1: No direct costs; Study 2: Direct cost= $118.17, cost-per-participant= $1.94

XSL•FO
RenderX

Participant Engagement and Retention

The amount of time participants spend using the site or app in a study can affect retention rates [37], and retention can greatly impact the overall power and generalizability of study results [22]. Engagement is a complex amalgam of time and interactivity with the website or app [37]. Examples of engagement metrics include number of times logging in to a website, number of page or icon clicks within an app, or completing surveys [38]. Slightly more than half (58%, 7/12) of the studies were conducting interventional research, while five (42%) of the 12 studies were conducting survey research. The studies conducting survey research (42%, 5/12) reported participant engagement as the time it took to complete study surveys. Of the remaining articles (58%, 7/12), only three (43%) reported participant engagement defined as the amount of time participants spent using the site. One intervention entailed an email chain lasting for 6 weeks [27]. Another study required a 90-day Facebook intervention and periodic biochemical confirmation of smoking cessation for 3 months [31]. The third study involved random assignment to one of two online interventions, each lasting 3 months before follow-up [29]. A third (4/12) of all the articles did not specify participant engagement at all.

The studies conducting survey research used survey completion as participant engagement. These five studies reported the number of participants that began eligibility screening, were deemed eligible, and successfully completed the surveys [25,30,32,33,36]. From these data, a completion rate was calculated as a ratio of the number of participants who completed surveys over eligible participants and then reported. These five articles (42%) either did not require or report survey follow-up, so their retention rates were calculated from their completion rates. However, retention in the context of mHealth and Web-based studies is focused on retaining participants for longer durations of time, so these rates were not reported.

Only one study (8%) reported participant retention [29]. They specified retention rates by recruitment method and found the following rates over a 3-month period: overall (52%), standard media (53%), broadcast email (46%), word of mouth (62%), press releases on health websites (56%), social media (paid ads and free posts—64%), and Google Ad Words (46%) (Tables 1 and 2). This study found no significant difference in retention rates between Web-based methods and traditional methods of recruitment [29].

XSL•FO
RenderX

Table 2. Summary of studies—Intervention, results, and retention.

Authors (year)	Intervention	Intervention duration	Results by recruitment method	Retention methods	Retention or completion rates
Batterham (2014)	Online survey	Time taken to complete online survey	Online surveys cost less than postal surveys- Internal links more cost less than external links. Content of ads crucial to the cost aspect of online recruitment. Terms "mental health problems" was more effective than "emotional well-being."	Completion rate percentage calculated in order to see how many people completed surveys after responding to ads.	Survey: 10.4% completion for Problem/Altruistic ad; 11% completion for Problem/Self-Gain ad; 5.8% completion for Positive/Altruistic ad; 9% completion for Positive/Self-Gain ad
Fenner (2012)	Health survey	15-30 minutes to complete survey, either electronically or at the site	Facebook recruitment found to compare favorably with traditional recruitment methods. Facebook also found to yield a representative sample	Calculated a participation rate for those who completed the survey out of those who clicked on the ad	Survey: 3.5% participation rate out of all who clicked on ad, 65% survey completion of those who consented
Frandsen, Walters, & Ferguson (2013)	[To be reported in future publication]	[To be reported in future publication]	Most participants recruited through online methods (Facebook), Facebook cost twice as much per participant than print media. Participant demographics from each method of recruitment were equally matched meaning online methods can supplement traditional methods.	Not reported	Not reported
Graham, Bock, Cobb, Niaura, & Abrams (2006)	Telephone counseling or using QuitNet.com (an Internet smoking cessation website)	N/A	Google yielded the greatest number of participants	Completing baseline assessment	51.3% completed baseline assessment and were randomized to treatment
Graham, Milner, Saul, & Pfaff (2008)	Telephone counseling and using QuitNet.com for smoking cessation treatment	[Not reported]	Traditional methods yielded more participants and found those participants engaged with the website more, online methods reached typically hard-to-reach populations and was found to cost less	Number of logins, minutes per login, number of page views, and interactions with other users and counselors	18.4% of identifiers on QuitNet were from traditional media, 81.6% from online media. 9.1% of online clicks registered for cessation treatment (6.8% Web-only, 1.1% phone only, 1.2% both); retention data not available at time of analysis
Heffner, Wyszynski Comstock, Mercer, & Bricker (2013)	Baseline survey, 3-month follow-up assessment	Time taken to complete baseline survey and follow-up 3 months later	GoogleAds had highest participant yield, social media and GoogleAds cost more than traditional methods. No difference between traditional and online methods in data retention or intervention success	Completing follow-up after 3 months went by	3-month retention: Overall: 52%, Standard media: 53%, email: 46%, word of mouth: 62%, medical Internet: 56%, social media: 64%, Google: 46%
Morgan, Jorm, & Mackinson (2013)	Patient Health Questionnaire, receiving weekly emails with self-help strategies	6 weeks	Google had highest participant yield, found to be less time-consuming and more effective than other recruitment techniques, even those that are free.	Completing baseline survey, receiving emails, and completing depression questionnaire	Survey: 78% completion after consent
Ramo, Hall, & Prochaska (2010)	10-item smoking questionnaire; Fagerstron Test of Nicotine Dependence; smoking stages of change questionnaire	20 min survey	Adbrite Internet advertisements resulted in highest participant yield (Facebook was most successful Internet Web sites), craigslist and SSI were more successful at targeting young adult smokers	Completing survey in entirety (all questionnaires)	Survey: 59.8% completion after eligibility screening

XSL•FO
RenderX

Authors (year)	Intervention	Intervention duration	Results by recruitment method	Retention methods	Retention or completion rates
Ramo & Prochaska (2012)	Smoking Stages of Change Questionnaire, A Smoking Questionnaire, Thoughts about Abstinence form	Time to complete survey	Facebook found to be more affordable than other previously-used methods with a similar population (Internet marketing company cost-per-participant=$42.77), success of specific ads dependent upon Facebook's approval of the ad	Completing survey in entirety	Survey: 50% completion after eligibility screening
Ramo, Rodriguez, Chavez, Sommer, & Prochaska (2014)	Smoking History Questionnaire, Smoking States of Change scale, baseline survey, Private Facebook group tailored to their readiness to quit (ready, thinking about it, not ready), bioconfirmation of nonsmoking tests	One year	Facebook found to be efficient and affordable. Newsfeed ads more successful—Ads with simple pictures of cigarettes, the study logo, and general info messaging were most successful	Participate in a Facebook group and complete follow-up assessments with saliva cotinine tests at 3, 6, and 12 months.	34% of those who were eligible and consented completed the intervention
Raviotta, Nowalk, Lin, Huang, & Zimmerman (2014)	Complete baseline survey and postvaccination survey. Half participants randomized to standard dosing group (0, 2, & 6 months) and half randomized to alternate dosing group (0, 2, & 12 months).	[Not reported]	Traditional methods reached more people, but online methods were more likely to reach hard-to-reach populations. Direct cost was higher for electronic advertising	Completing four visits to study site for intake survey, blood draw, vaccination does 1, 2, & 3, and final blood draw and survey	Survey: 70.7% completion after eligibility screening
Valdez, Guterbock, Thompson, Reilly, Menefee, Bennici, Williams, & Rexrode (2014)	Survey	Time taken to complete the survey	Facebook found to be affordable, but not feasible for large, quantitative studies	Completing the survey in its entirety	Survey: 77.2% completion after eligibility screening

Discussion

Principal Findings

The use of mHealth has tremendous potential for improving public health through its convenience, wide reach, and flexibility [14]. The Internet and mHealth apps are increasingly used by individuals who seek health information, thus increasing the potential to reach underserved populations [32]. However, there is a need to develop and evaluate mHealth apps in order to establish consistent, effective methods for producing health behavior change. Research on mHealth may experience challenges in participant recruitment and retention due to its very nature, which is characterized by the virtual aspect of the intervention [27]. Virtual interventions may lead to less investment and fraudulent enrollment on behalf of participants due to lack of relationship with the study team and potential incentives to participate [27]. Personal relationships may help participants to understand their contribution to the research [27]. Currently the literature on Web-based interventions is greater than that for mHealth apps, but there is a lack of detail in both about participant recruitment and even less information on retention. More attention must be paid to these two factors as they affect the validity and generalizability of the research findings [15].

Cost of Recruitment

For researchers conducting online and/or mHealth studies, online methods of recruitment have the potential to achieve better affordability than traditional methods of recruitment [27]. However, the results of this review found inconsistent findings related to cost-per-participant. This review also found conflicting outcomes regarding whether Web-based or traditional methods of recruitment were more effective at enrolling participants. It appears that the type of intervention and target population influenced which type of recruitment method was most successful.

Of the 12 studies examined, there was no clear consensus on which method (Web or traditional) is best to use when conducting mHealth research. The success of recruitment method varied widely based on population, budget, intervention, cost, and study design. "Best method", as defined in these articles, was the one that achieved the highest participant yield for the cost incurred. Assessing largest participant yield can be achieved through analyzing overall impressions and click-through rates. Four articles (33%) found Facebook ads to be effective for recruitment when the intention was to achieve a sample within a specific age range. The use of Facebook ads for recruiting those under 25 may be due to the fact that Facebook offers the ability to target a specific demographic or age group by displaying ads only on profiles within that listed age range. Google ads appeared to be the most effective

XSL•FO
RenderX

Web-based recruitment method when there was no specific age range targeted. Four studies (33%) compared Web-based methods against traditional methods of recruitment and reached different conclusions about highest participant yield and affordability.

There are a variety of factors that could lead to these results. First, paying for advertisements is different for each online method. For example, the bidding method that both Facebook and Google use does not allow for a stable cost prediction that could be replicated by the same researcher or even future researchers. This is due to the intermittent presence of other bidders. Based on project budgets and desired participant demographics, very different results could be achieved. In addition, the business models of Internet advertising (eg, Facebook and Google ads) are constantly changing, presenting ongoing challenges for determining affordable strategies relating to online recruitment. Also, virtual interventions and methods of recruitment allow for a certain degree of ambiguity and anonymity. This might allow participants to feel more comfortable enrolling in a study or could lead to less investment in the process.

Participant Retention

Overall, studies are not doing a good job of reporting retention rates. Only one of the five studies that utilized a long-duration intervention reported participant retention rates. Retention issues can create bias and complications with generalizability [22], and if the rates are not reported, it is unclear if the results are valid. Consort diagrams are needed to describe the recruitment process and sources of attrition. Retention is a factor that not only should be addressed, but also reported with its potential effects in all articles related to online (eHealth) and mHealth interventions.

We recommend that researchers report recruitment methods, costs, participant demographics, and participant engagement metrics in their publications. Studies evaluating the effectiveness of interventions, especially mHealth interventions, should include this information in order to allow for future replication. A set of guidelines should be established for informing eHealth and mHealth researchers about recruiting generalizable samples. These guidelines might include, for example, a list of successful methods for targeting a specific participant demographic or approaches to maximizing a limited recruitment budget.

Limitations

This review has two possible limitations. First, the limited number of studies required us to modify the selection process to more broadly examine the literature in order to meet the specific needs of the research team. Therefore, our inclusionary and exclusionary criteria changed over the course of the study. Second, the studies identified during the search used very different designs and methods. These discrepancies made it difficult to form recommendations as to which methods of recruitment are most beneficial and cost-effective in conducting mHealth research.

Future Research and Recommendations

Further evaluation of online methods of recruitment is necessary to better understand their effectiveness for use in Web-based and mHealth research. Due to the unique challenges that Web-based and mHealth interventions face, there is a critical need for published articles focusing solely on recruitment methods, including participant recruitment and retention rates. Although it was outside the scope of this review, researchers in this area may also experience the possibility of fraudulent participants. Little is known about how the mHealth research process may encourage/discourage “fake” participants from enrolling in this type of study. There is also a need for guidelines and recommendations for affordable ways of recruiting and retaining representative samples of participants in Web-based and mHealth research. It is difficult to assess which methods of recruitment will work best, and there is not a “one-size-fits-all” approach, which should be included as a part of a set of guidelines. Currently, recommendations for guidelines cannot be made due to the inconsistent nature of reporting of online recruitment strategies. As the body of literature grows around mHealth, and as researchers begin to report specific recruitment methods, associated costs, and retention methods and rates, guidelines may be formed.

Conclusions

Research on mHealth apps and Web-based interventions may experience challenges with participant recruitment and retention. The research available regarding ideal participant yield and the associated costs of online recruitment methods for Web-based and mHealth studies is minimal and inconsistent. Researchers in this area should routinely report metrics regarding recruitment methods used and participant attrition. This includes detailing specifically what type of ads were used (eg, banner or search ads), as well as specific cost information (eg, cost per click, cost per participant, and overall costs). In addition, retention methods and retention rates should be included, as it is an important factor for mHealth and Internet-based studies. Information needed would include participant engagement in the intervention (eg, metrics associated with use of the intervention), and how many participants successfully completed the intervention and associated research activities. A set of guidelines for successful and affordable methods of recruitment, and metrics for evaluating mHealth engagement and participant retention are needed.

Conflicts of Interest

None declared.

References

XSL•FO
RenderX

1. Molla MT, Centers for Disease Control and Prevention (CDC). Expected years of life free of chronic condition-Induced activity limitations - United States, 1999-2008. MMWR Surveill Summ 2013 Nov 22;62:Suppl 3:87-Suppl 3:91 [FREE Full text]
2. Manca DP, Aubrey-Bassler K, Kandola K, Aguilar C, Campbell-Scherer D, Sopcak N, et al. Implementing and evaluating a program to facilitate chronic disease prevention and screening in primary care: a mixed methods program evaluation. Implement Sci 2014;9:135 [FREE Full text] [doi: 10.1186/s13012-014-0135-7] [Medline: 25293785]
3. Ventola CL. Mobile devices and apps for health care professionals: uses and benefits. P T 2014 May;39(5):356-364 [FREE Full text] [Medline: 24883008]
4. Becker S, Miron-Shatz T, Schumacher N, Krocza J, Diamantidis C, Albrecht U. mHealth 2.0: Experiences, Possibilities, and Perspectives. JMIR Mhealth Uhealth 2014;2(2):e24 [FREE Full text] [doi: 10.2196/mhealth.3328] [Medline: 25099752]
5. Pew Research Internet Project. 2014 Jan. Device Ownership Over Time URL: http://www.pewinternet.org/data-trend/mobile/device-ownership/ [accessed 2015-02-10] [WebCite Cache ID 6WFCT9Iaz]
6. Fox S, Duggan M. Pew Internet Project. Mobile Health 2012 URL: http://www.pewinternet.org/files/old-media//Files/Reports/2012/PIP_MobileHealth2012_FINAL.pdf [accessed 2015-06-19] [WebCite Cache ID 6ZPIXzo6V]
7. Duggan M. Pew Internet Project. Cell Phone Activities 2013 URL: http://www.pewinternet.org/2013/09/19/cell-phone-activities-2013/ [accessed 2015-02-10] [WebCite Cache ID 6WFCdumGH]
8. Fox S, Duggan M. Pew Research Center. Health Online 2013 URL: http://www.pewinternet.org/files/old-media//Files/Reports/PIP_HealthOnline.pdf [accessed 2015-06-19] [WebCite Cache ID 6ZPIpRNTh]
9. Purcell K, Rainie L. Pew Research Center. Americans Feel Better Informed Thanks to the Internet URL: http://www.pewinternet.org/files/2014/12/PI_InformedWeb_120814_02.pdf [accessed 2015-06-19] [WebCite Cache ID 6ZPIwe1We]
10. Charani E, Castro-Sánchez E, Moore LSP, Holmes A. Do smartphone applications in healthcare require a governance and legal framework? It depends on the application!. BMC Med 2014;12:29 [FREE Full text] [doi: 10.1186/1741-7015-12-29] [Medline: 24524344]
11. Wang A, An N, Lu X, Chen H, Li C, Levkoff S. A classification scheme for analyzing mobile apps used to prevent and manage disease in late life. JMIR Mhealth Uhealth 2014;2(1):e6 [FREE Full text] [doi: 10.2196/mhealth.2877] [Medline: 25098687]
12. Armstrong KA, Semple JL, Coyte PC. Replacing ambulatory surgical follow-up visits with mobile app home monitoring: modeling cost-effective scenarios. J Med Internet Res 2014;16(9):e213 [FREE Full text] [doi: 10.2196/jmir.3528] [Medline: 25245774]
13. Danaher BG, McKay HG, Seeley JR. The information architecture of behavior change websites. J Med Internet Res 2005;7(2):e12 [FREE Full text] [doi: 10.2196/jmir.7.2.e12] [Medline: 15914459]
14. Middelweerd A, Mollee JS, van der Wal CN, Brug J, Te Velde SJ. Apps to promote physical activity among adults: a review and content analysis. Int J Behav Nutr Phys Act 2014;11:97 [FREE Full text] [doi: 10.1186/s12966-014-0097-9] [Medline: 25059981]
15. van Velthoven MH, Li Y, Wang W, Du X, Chen L, Wu Q, et al. mHealth Series: Factors influencing sample size calculations for mHealth-based studies - A mixed methods study in rural China. J Glob Health 2013 Dec;3(2):020404 [FREE Full text] [doi: 10.7189/jogh.03.020404] [Medline: 24363922]
16. Bricker J, Wyszynski C, Comstock B, Heffner JL. Pilot randomized controlled trial of web-based acceptance and commitment therapy for smoking cessation. Nicotine Tob Res 2013 Oct;15(10):1756-1764 [FREE Full text] [doi: 10.1093/ntr/ntt056] [Medline: 23703730]
17. Abroms LC, Lee WJ, Bontemps-Jones J, Ramani R, Mellerson J. A content analysis of popular smartphone apps for smoking cessation. Am J Prev Med 2013 Dec;45(6):732-736 [FREE Full text] [doi: 10.1016/j.amepre.2013.07.008] [Medline: 24237915]
18. Abroms LC, Padmanabhan N, Thaweethai L, Phillips T. iPhone apps for smoking cessation: a content analysis. Am J Prev Med 2011 Mar;40(3):279-285 [FREE Full text] [doi: 10.1016/j.amepre.2010.10.032] [Medline: 21335258]
19. Backinger CL, Augustson EM. Where there's an app, there's a way? Am J Prev Med 2011 Mar;40(3):390-391 [FREE Full text] [doi: 10.1016/j.amepre.2010.11.014] [Medline: 21335276]
20. Sully BGO, Julious SA, Nicholl J. A reinvestigation of recruitment to randomised, controlled, multicenter trials: a review of trials funded by two UK funding agencies. Trials 2013;14:166 [FREE Full text] [doi: 10.1186/1745-6215-14-166] [Medline: 23758961]
21. Ferwerda M, van Beugan S, van Burik A, van Middendorp H, de Jong EMGJ, van de Kerkhof PCM, et al. What patients think about E-health: patients' perspective on internet-based cognitive behavioral treatment for patients with rheumatoid arthritis and psoriasis. Clin Rheumatol 2013 Jun;32(6):869-873. [doi: 10.1007/s10067-013-2175-9] [Medline: 23354514]
22. Fewtrell MS, Kennedy K, Singhal A, Martin RM, Ness A, Hadders-Algra M, et al. How much loss to follow-up is acceptable in long-term randomised trials and prospective studies? Arch Dis Child 2008 Jun;93(6):458-461. [doi: 10.1136/adc.2007.127316] [Medline: 18495909]
23. Koo M, Skinner H. Challenges of internet recruitment: a case study with disappointing results. J Med Internet Res 2005;7(1):e6 [FREE Full text] [doi: 10.2196/jmir.7.1.e6] [Medline: 15829478]

XSL•FO
RenderX

24. Frandsen M, Walters J, Ferguson SG. Exploring the viability of using online social media advertising as a recruitment method for smoking cessation clinical trials. Nicotine Tob Res 2014 Feb;16(2):247-251. [doi: 10.1093/ntr/ntt157] [Medline: 24127266]
25. Fenner Y, Garland SM, Moore EE, Jayasinghe Y, Fletcher A, Tabrizi SN, et al. Web-based recruiting for health research using a social networking site: an exploratory study. J Med Internet Res 2012;14(1):e20 [FREE Full text] [doi: 10.2196/jmir.1978] [Medline: 22297093]
26. Graham AL, Milner P, Saul JE, Pfaff L. Online advertising as a public health and recruitment tool: comparison of different media campaigns to increase demand for smoking cessation interventions. J Med Internet Res 2008;10(5):e50 [FREE Full text] [doi: 10.2196/jmir.1001] [Medline: 19073542]
27. Morgan AJ, Jorm AF, Mackinnon AJ. Internet-based recruitment to a depression prevention intervention: lessons from the Mood Memos study. J Med Internet Res 2013;15(2):e31 [FREE Full text] [doi: 10.2196/jmir.2262] [Medline: 23403043]
28. Seltzer ED, Stolley MR, Mensah EK, Sharp LK. Social networking site usage among childhood cancer survivors--a potential tool for research recruitment? J Cancer Surviv 2014 Sep;8(3):349-354. [doi: 10.1007/s11764-014-0348-4] [Medline: 24532046]
29. Heffner JL, Wyszynski CM, Comstock B, Mercer LD, Bricker J. Overcoming recruitment challenges of web-based interventions for tobacco use: the case of web-based acceptance and commitment therapy for smoking cessation. Addict Behav 2013 Oct;38(10):2473-2476 [FREE Full text] [doi: 10.1016/j.addbeh.2013.05.004] [Medline: 23770645]
30. Batterham PJ. Recruitment of mental health survey participants using Internet advertising: content, characteristics and cost effectiveness. Int J Methods Psychiatr Res 2014 Jun;23(2):184-191. [doi: 10.1002/mpr.1421] [Medline: 24615785]
31. Ramo DE, Rodriguez TMS, Chavez K, Sommer MJ, Prochaska JJ. Facebook Recruitment of Young Adult Smokers for a Cessation Trial: Methods, Metrics, and Lessons Learned. Internet Interv 2014 Apr;1(2):58-64 [FREE Full text] [doi: 10.1016/j.invent.2014.05.001] [Medline: 25045624]
32. 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]
33. Ramo DE, Hall SM, Prochaska JJ. Reaching young adult smokers through the internet: comparison of three recruitment mechanisms. Nicotine Tob Res 2010 Jul;12(7):768-775 [FREE Full text] [doi: 10.1093/ntr/ntq086] [Medline: 20530194]
34. Graham AL, Bock BC, Cobb NK, Niaura R, Abrams DB. Characteristics of smokers reached and recruited to an internet smoking cessation trial: a case of denominators. Nicotine Tob Res 2006 Dec;8 Suppl 1:S43-S48 [FREE Full text] [Medline: 17491170]
35. Raviotta JM, Nowalk MP, Lin CJ, Huang H, Zimmerman RK. Using Facebook™ to Recruit College-Age Men for a Human Papillomavirus Vaccine Trial. Am J Mens Health 2014 Nov 11:-. [doi: 10.1177/1557988314557563] [Medline: 25389213]
36. Valdez RS, Guterbock TM, Thompson MJ, Reilly JD, Menefee HK, Bennici MS, et al. Beyond traditional advertisements: leveraging Facebook's social structures for research recruitment. J Med Internet Res 2014;16(10):e243 [FREE Full text] [doi: 10.2196/jmir.3786] [Medline: 25348050]
37. Danaher BG, Boles SM, Akers L, Gordon JS, Severson HH. Defining participant exposure measures in Web-based health behavior change programs. J Med Internet Res 2006;8(3):e15 [FREE Full text] [doi: 10.2196/jmir.8.3.e15] [Medline: 16954125]
38. Couper MP, Alexander GL, Zhang N, Little RJA, Maddy N, Nowak MA, et al. Engagement and retention: measuring breadth and depth of participant use of an online intervention. J Med Internet Res 2010;12(4):e52 [FREE Full text] [doi: 10.2196/jmir.1430] [Medline: 21087922]

Edited by G Eysenbach; submitted 19.02.15; peer-reviewed by M Hingle, E Augustson; comments to author 02.04.15; revised version received 14.05.15; accepted 13.06.15; published 22.07.15.

Please cite as:
Lane TS, Armin J, Gordon JS
Online Recruitment Methods for Web-Based and Mobile Health Studies: A Review of the Literature
J Med Internet Res 2015;17(7):e183
URL: http://www.jmir.org/2015/7/e183/
doi:10.2196/jmir.4359
PMID:26202991

©Taylor S Lane, Julie Armin, Judith S Gordon. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 22.07.2015. 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

Allowing Physicians to Choose the Value of Compensation for Participation in a Web-Based Survey: Randomized Controlled Trial

Alison E Turnbull[1,2,3], MPH, DVM, PhD; Cristi L O'Connor[2], MHS; Bryan Lau[2], ScM, MHS, PhD; Scott D Halpern[4,5], MD, PhD; Dale M Needham[1,3,6], MD, PhD, FCPA

[1]School of Medicine, Division of Pulmonary and Critical Care Medicine, Johns Hopkins University, Baltimore, MD, United States
[2]Bloomberg School of Public Health, Department of Epidemiology, Johns Hopkins University, Baltimore, MD, United States
[3]Outcomes After Critical Illness and Surgery Group, Johns Hopkins University, Baltimore, MD, United States
[4]Perelman School of Medicine, Division of Pulmonary, Allergy and Critical Care, University of Pennsylvania, Philadelphia, PA, United States
[5]Perelman School of Medicine, Department of Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, PA, United States
[6]School of Medicine, Department of Physical Medicine and Rehabilitation, Johns Hopkins University, Baltimore, MD, United States

Corresponding Author:
Alison E Turnbull, MPH, DVM, PhD
School of Medicine
Division of Pulmonary and Critical Care Medicine
Johns Hopkins University
1830 E Monument St.
5th Floor
Baltimore, MD, 21205
United States
Phone: 1 4109552190
Fax: 1 4109550036
Email: turnbull@jhmi.edu

Abstract

Background: Survey response rates among physicians are declining, and determining an appropriate level of compensation to motivate participation poses a major challenge.

Objective: To estimate the effect of permitting intensive care physicians to select their preferred level of compensation for completing a short Web-based survey on physician (1) response rate, (2) survey completion rate, (3) time to response, and (4) time spent completing the survey.

Methods: A total of 1850 US intensivists from an existing database were randomized to receive a survey invitation email with or without an Amazon.com incentive available to the first 100 respondents. The incentive could be instantly redeemed for an amount chosen by the respondent, up to a maximum of US $50.

Results: The overall response rate was 35.90% (630/1755). Among the 35.4% (111/314) of eligible participants choosing the incentive, 80.2% (89/111) selected the maximum value. Among intensivists offered an incentive, the response was 6.0% higher (95% CI 1.5-10.5, *P*=.01), survey completion was marginally greater (807/859, 94.0% vs 892/991, 90.0%; *P*=.06), and the median number of days to survey response was shorter (0.8, interquartile range [IQR] 0.2-14.4 vs 6.6, IQR 0.3-22.3; *P*=.001), with no difference in time spent completing the survey.

Conclusions: Permitting intensive care physicians to determine compensation level for completing a short Web-based survey modestly increased response rate and substantially decreased response time without decreasing the time spent on survey completion.

(J Med Internet Res 2015;17(7):e189) doi:10.2196/jmir.3898

KEYWORDS

data collection; monetary incentives; cash; physicians; electronic questionnaire; survey design; response rate

XSL•FO
RenderX

Introduction

Understanding the opinions and practices of health care providers is essential for clinical research [1]. However, surveys of health care providers are plagued by declining response rates [2-4], and techniques for increasing response rates in the general public [5] frequently fail to motivate physicians [6]. Although low response rates do not necessarily bias results [7,8], they do increase the potential for nonresponse bias, and hamper publication [2,9-11].

While physician response rates are declining, inexpensive tools for conducting sophisticated Web-based electronic surveys are flourishing. Although physician response rates to electronic surveys have generally been lower than to postal surveys, many trials comparing postal versus electronic surveys were conducted a decade ago and targeted community-based physicians in regions where high-speed Internet access was unreliable or nonexistent [12-15]. As high-speed Internet access becomes ubiquitous and health care providers become more comfortable with Web-based technologies, electronic surveys have the potential to provide researchers with unique tools, including instant compensation for participation and data about how physicians interact with surveys.

A major challenge when designing a survey is determining the appropriate level of financial compensation required to incentivize participation [16-18]. Allowing respondents to choose how much they wish to be compensated provides insight into participant motivation and may maximize the cost-effectiveness of incentives by not spending funds on participants who do not require a large incentive to respond. An additional challenge posed by electronic surveys is the inability to provide prepayment. Prepayment in postal surveys, traditionally achieved by including cash in the survey envelope, is associated with significantly greater response rates among surveys of physicians versus providing payment contingent on survey completion [16,19,20].

We combined three techniques to address these challenges. First, we invited physicians to select their preferred level of instant compensation, up to US $50, for completing a short, electronic survey. Second, we attempted to engender altruism by reminding physicians that the study was funded by a limited student budget. Finally, compensation was only promised to the first 100 respondents, making it a scarce, time-limited incentive. To assess whether these three combined techniques affected response rate, time to response, survey completion rate, and time spent completing the electronic survey, we designed a randomized controlled trial of respondent-selected compensation.

Methods

Study Design

A previously described database of academic intensivists was used to recruit faculty from US hospitals with training programs accredited by the Accreditation Council for Graduate Medical Education in Internal Medicine-Critical Care Medicine, Anesthesiology-Critical Care Medicine, and Surgical Critical Care [21]. The database was updated in 2012 to include demographic and electronic contact information for 2482 physicians. Physicians were excluded from randomization if they (1) lacked electronic contact information (285/2482, 11.48%), (2) had been invited to participate in a pilot study (268/2482, 10.80%), (3) had made a previous request not to be contacted (76/2482, 3.06%), or (4) contributed to study design or survey development (3/2482, 0.12%). The remaining 1850 intensivists were potentially eligible to participate in a randomized trial of an intervention to increase communication about life support with families of critically ill patients [22], administered using the Qualtrics Web-based survey platform [23]. The Institutional Review Board (IRB) of Johns Hopkins School of Medicine approved the study. Intensivists were notified that survey completion served as consent to participate in the trial.

Randomization was blocked on intensivist sex, specialty—medicine, anesthesiology, or surgery—years since completing residency, and geographic region of residency [24,25]. Within each block, 45% of eligible intensivists were randomly assigned to the group with the ability to select their preferred level of compensation as an incentive to participate.

On November 20, 2012, each randomized intensivist was sent an invitation by email containing a unique link to the survey. All invitations included the survey topic, number of questions, expected time required to complete the survey (5 minutes), IRB approval, study confidentiality, number of follow-up/reminder emails for nonresponders, planned date for study closing (December 20, 2012), and names and affiliations of study investigators. Invitations for intensivists randomized to receive an incentive to participate also included the following text:

> *In appreciation for your participation, the first 100 respondents to complete the survey will be offered an Amazon.com gift code at the end of the survey. The code can be redeemed immediately for any amount up to $50. In selecting the compensation amount, please consider that this is a PhD thesis project being funded by a limited student budget.*

Reminder emails were sent to all intensivists who had not completed the survey on days 13, 22, and 28 following the initial invitation. In each of the reminder emails, intensivists randomized to the incentive group were informed that funds for incentives were still available. Because not all respondents chose to take the full amount available, there were sufficient funds to offer the incentive to more than 100 respondents. To establish survey eligibility, participants were first asked if they had treated patients in the intensive care unit (ICU) setting during the previous 2 years. Those who had were then asked one question about practice history followed by 10 screens, each containing a brief clinical scenario for review.

Participants randomized to the incentive intervention who completed the survey had the option of entering the amount they wished to spend at Amazon.com [26] up to US $50 using the Amazon Gift Codes On Demand service, which allows study participants to claim incentives instantly. Among those participants who elected to receive one, incentives were claimed on the Amazon.com website. The value of created incentives

XSL•FO
RenderX

was instantly deducted from a study fund containing US $5000. Study investigators could not access information on goods purchased by participants or the timing of purchases made using gift codes.

Statistical Analysis

The primary outcome measure was the difference in survey response rate between trial arms. Response rates were calculated in accordance with response rate three (RR3), defined by the American Association for Public Opinion Research guidelines for Internet surveys of specifically named persons [27] as follows:

RR3 = I/(I + P + *eC*[UH + UO] + *eI*[UH + UO])

The RR3 is equal to the number of eligible participants who responded to all survey questions (I) divided by the sum of eligible participants who responded to all survey questions (I), eligible participants who answered the eligibility screening question but did not answer all survey questions (P), and the estimated proportion in the control arm (*eC)* and intervention arm (*eI)* of nonresponders (UH) and responders with unknown eligibility (UO) who were eligible. The proportions of eligible nonresponders and responders with unknown eligibility were estimated based on the proportion of responders in each arm of the trial who answered the screening question and were known to be eligible.

Secondary outcome measures were defined as follows: survey completion rate (ie, number of intensivists known to be eligible who answered all survey questions divided by the number of eligible intensivists who clicked the link to the Web-based survey), time to response (ie, time of survey completion minus the time the initial email was sent among intensivists who completed the survey), and time spent completing survey (ie, time of survey completion minus time that the link in the invitation email was clicked by eligible intensivists).

Analyses were performed using the R programming language version 3.0.1 (Vienna, Austria) [28] using two-sided significance tests, with $P<.05$ indicating statistical significance, and data were analyzed on an intention-to-treat basis. Hypothesis tests for differences in proportion were performed using Pearson's chi-square test. Fisher's exact test was used when a cell within a contingency table contained fewer than 10 observations. Confidence intervals for differences in proportions were calculated using the Wald interval. Differences in the distribution of continuous variables were assessed using the Wilcoxon-Mann-Whitney test [29].

Results

Overview

The overall response rate was 35.90% (630/1755), with 92.0% (630/685) of eligible respondents answering all questions in the survey (see Figure 1 and Table 1). A total of 13 out of 991 (1.3%) respondents in the control arm and 22 out of 859 (2.6%) respondents in the incentive arm indicated that they had not treated patients in the ICU setting in the last 2 years and were deemed ineligible. Among the 55 known eligible respondents who did not complete the survey, 32 (58%) answered the screening question and the question about practice history, but did not answer any of the questions related to the brief clinical scenarios. The remaining 23 (42%) participants responded to a median of 4 scenarios (interquartile range [IQR] 3-6). The median time to response was 3.4 days (IQR 0.3-22.0) and the median amount of time eligible intensivists spent completing the survey was 3.9 minutes (IQR 2.5-5.5). Among eligible intensivists invited to participate, 80.49% (1489/1850) were male, 63.30% (1171/1850) specialized in internal medicine, and the median number of years since completing initial residency was 20 (IQR 13-28). The characteristics of intensivists randomized to the control versus incentive groups were similar (see Table 2).

Based on the RR3 equation, the overall response rate was equal to 35.90%:

(316+314)/([316+314]+[35+20]+ .96 [624+3]+.93[500+3])= 630/1755 = 35.90%

The control arm response rate was equal to 33.2%:

316/(316+35+ .96 [624+3])= 316/953=33.2%

The incentive arm response rate was equal to 39.2%:

314/(314+20+ .93[500+3])= 314/802=39.2%

XSL•FO
RenderX

Table 1. Response rate calculation values using response rate three (RR3)[a].

Term	Definition	Control arm (n=991), n (%) or n/n (proportion)	Incentive arm (n=859), n (%) or n/n (proportion)
I	Eligible participants who answered all survey questions, n (%)	316 (31.9)	314 (36.6)
P	Participants who answered the eligibility question but did not answer all survey questions, n (%)	35 (3.5)	20 (2.3)
UH	Nonresponders, n (%)	624 (63.0)	500 (58.2)
UO	Responders with unknown eligibility, n (%)	3 (0.3)	3 (0.3)
eC	Estimated proportion of eligible participants in the control arm, n/n (proportion)	351/367 (0.96)	N/A[b]
eI	Estimated proportion of eligible participants in the incentive arm, n/n (proportion)	N/A	334/359 (0.93)

[a]RR3 = I/(I + P + *eC*[UH + UO] + *eI*[UH + UO]). RR3 is defined by the American Association for Public Opinion Research guidelines for Internet surveys of specifically named persons [27].

[b]Not applicable (N/A).

Table 2. Characteristics of the study population[a].

Variable	Control (n=991), n (%) or median (IQR[b])	Incentive (n=859), n (%) or median (IQR)
Male, n (%)	786 (79.3)	703 (81.8)
Specialty, n (%)		
Medicine	615 (62.1)	556 (64.7)
Surgery	204 (20.6)	170 (19.8)
Anesthesia	172 (17.4)	133 (15.5)
Years since residency, median (IQR)	20 (13-28)	20 (13-27)
Years since residency not reported, n (%)	156 (15.7)	134 (15.6)
Region of residency [c], n (%)		
Northeast	314 (32.0)	291 (33.9)
Midwest	215 (21.7)	189 (22.0)
South	218 (22.0)	192 (22.4)
West	113 (11.4)	97 (11.3)
International	55 (5.5)	36 (4.2)
Unknown	73 (7.4)	54 (6.3)

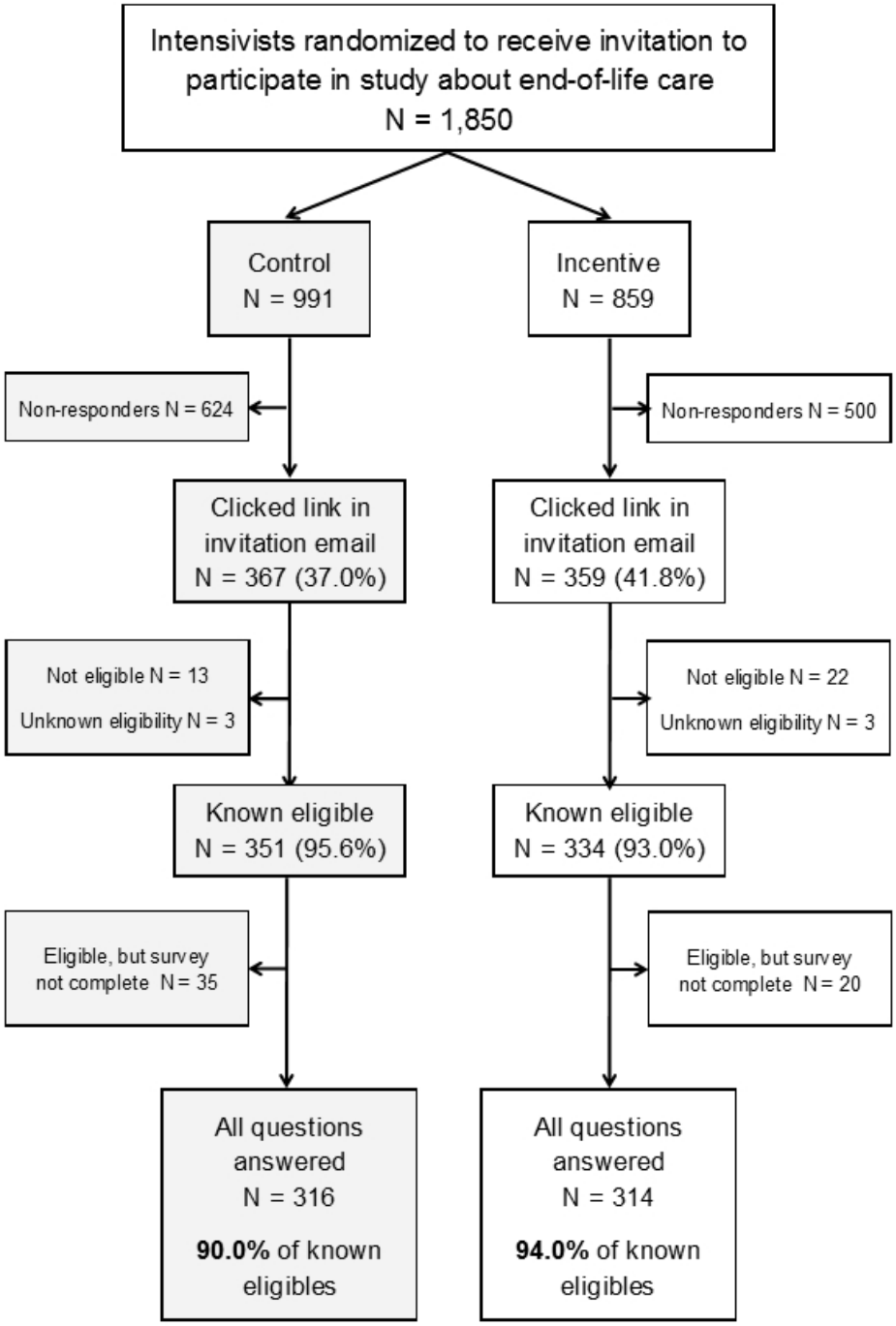
[a]Percentages may not sum to 100% due to rounding.

[b]Interquartile range (IQR).

[c]Region defined according to US census region.

XSL•FO
RenderX

Figure 1. Study profile. Eligible participants had treated patients in the intensive care unit (ICU) setting during the last 2 years.



Primary and Secondary Outcomes

The overall response rate among eligible intensivists offered an incentive was 39.2% (314/802) versus 33.2% (316/953) in the control group (*P*=.01) (see Figure 1). The proportion of eligible respondents answering all survey questions was modestly greater in the incentive group (807/859, 94.0%) versus the control group (892/991, 90.0%) (*P*=.06). Among the known eligible participants who did not complete the survey, 65% (13/20) in the incentive group and 54% (19/35) in the control group did not answer questions related to the brief clinical scenarios (*P*=.44). In contrast to these relatively small effects on response, the incentive was associated with a large reduction in median time to response among responders (0.8 days for incentive group and 6.6 days for control group, *P*=.001; see Table 3). The median time required to complete the survey was 3.9 minutes in each group (*P*=.56).

XSL•FO
RenderX

Table 3. Survey outcomes by intervention arm.

Survey outcomes	Control (n=991), n (%) or median (IQR[a])	Incentive (n=859), n (%) or median (IQR)	*P* [b]
Response rate[c], n (%)	329 (33.2)	337 (39.2)	.01
Eligible responders completing survey, n (%)	892 (90.0)	807 (94.0)	.06
Days to response among eligible responders, median (IQR)	6.6 (0.3-22.3)	0.8 (0.2-14.4)	.001
Minutes spent completing survey, median (IQR)	3.9 (2.5-5.5)	3.9 (2.4-5.5)	.56

[a]Interquartile range (IQR).

[b]Calculated using the chi-square test for proportions and the Wilcoxon-Mann-Whitney test for continuous variables.

[c]Response rate calculated in accordance with response rate three (RR3) defined by the American Association for Public Opinion Research guidelines for Internet surveys of specifically named persons [27].

Incentive Amount

Out of 314 intensivists who answered all survey questions and were offered an incentive, 111 (35.4%) chose to create one (see Table 4). All who chose to accept an incentive were able to do so, as US $95 remained at study end from the US $5000 originally budgeted for incentives. Overall, 80.2% (89/111) of accepted incentives were for the maximum value of US $50. Among the 22 intensivists out of 111 (19.8%) who chose incentives worth less than US $50, the median value was US $20 (IQR $11-$25). Intensivists randomized to the incentive group who accepted, versus did not accept, incentives completed their first residency more recently (median years since residency 15 vs 19, respectively; *P*=.004). A greater proportion of male versus female intensivists chose incentives (99/259, 38.2% versus 12/55, 22%, respectively; *P*=.03).

Table 4. Participant characteristics and survey outcomes within the incentive arm of the trial, by participant response to financial incentive[a].

Variable	Incentive declined (n=203), n (%) or median (IQR[b])	Incentive claimed (n=111) , n (%) or median (IQR)	*P* [c]
Male, n (%)	160 (78.8)	99 (89.2)	.03
Specialty, n (%)			.17
Medicine	128 (63.1)	68 (61.3)	
Surgery	43 (21.2)	17 (15.3)	
Anesthesia	32 (15.8)	26 (23.4)	
Years since residency, median (IQR)	19 (13-27)	15 (10-22)	.004
Region of residency [d], n (%)			.63
Northeast	68 (33.5)	33 (29.7)	
Midwest	55 (27.1)	24 (21.6)	
South	44 (21.7)	30 (27.0)	
West	25 (12.3)	10 (9.0)	
International	6 (3.0)	4 (3.6)	
Unknown	5 (2.5)	10 (9.0)	
Days to response among responders, median (IQR)	0.5 (0.2-13.4)	2.9 (0.3-18.5)	.10
Minutes spent completing survey, median (IQR)	3.7 (2.4-5.5)	4.2 (2.6-5.5)	.66

[a]Percentages may not sum to 100% due to rounding.

[b]Interquartile range (IQR).

[c]Calculated using Fisher's exact test, the chi-square test for proportions, and the Wilcoxon-Mann-Whitney test for continuous variables.

[d]Region defined according to US census region.

XSL•FO
RenderX

Discussion

Principal Findings

In a national randomized trial of 1850 academic intensivists, permitting these physicians to choose their preferred level of financial compensation for participating in a short Web-based survey resulted in a 6.0% (95% CI 1.5-10.5, *P*=.01) absolute increase (15.3% relative increase) in response rate, a 3.9% absolute increase in survey completeness, and a faster response time (0.8 vs 6.6 days), with no impact on the time spent completing the survey. Although 66.7% of intensivists offered compensation did not take it, those who did accept it generally took the maximum US $50 amount that was available to them.

Among intensivists offered an incentive, the only respondent characteristics associated with taking it was time since completing residency and gender. More recent graduates of medical training are likely to have lower salaries, higher educational debt levels, and greater electronic expertise, making US $50 more valuable and accessible. Although previous studies have found male health care workers to be less likely to respond to surveys than women [2,10], a sex-based difference in response to compensation has not been commonly reported in prior literature and merits greater investigation. The observed association between instant compensation and time to response is likely to have been influenced by the perceived scarcity of the compensation (ie, the invitation email said the incentive would be offered to the first 100 respondents) and the proximity of the study timing to annual holiday spending in December. This enhanced desirability of scarce resources is a well-known psychological effect [30], and an example of the larger phenomenon of loss aversion [31].

Shortening the time to survey response decreases the number of reminder or follow-up contacts required. Sending fewer reminders saves time and money when surveys are administered by post. Additionally, previous work suggests that late responders often differ from early responders both demographically and in their survey responses [32-35]. Techniques that recruit physicians who intend to respond, but are prone to delaying participation, help ensure their unique perspectives are represented in the study sample.

The fact that relatively few intensivists (33.3%) took the incentive may have meant that most respondents were sufficiently interested in the survey topic not to require any further motivation for participation, or that the survey was short enough—median completion time was 3.8 minutes—that most respondents did not require reimbursement. Participants who took the incentive may have been less interested in the survey topic, but this difference in interest did not lead these participants to spend less time considering or answering survey questions. Efficiently incentivizing participants to thoughtfully answer all questions may be more important for lengthier research questionnaires, although recent trials have reported no association between survey length and physician response for either postal [36-38] or Web-based surveys [39]. Decisions about whether to take the incentive also may have been influenced by altruistic or sympathetic sentiments created by disclosing the limited funding available for this student thesis project.

As physician response rates decline, leveraging available funds to incentivize survey participation becomes increasingly important. Allowing physicians in the incentive group to decide whether they wished to be compensated and explicitly mentioning that only the first 100 respondents could take the incentive allowed us to offer the full US $50 to the 859 targeted intensivists who were potentially motivated by financial gain, despite a budget of only US $5000. There are two alternative incentive strategies to consider. The first is providing US $50 to all 314 eligible respondents who completed the survey. This would have cost US $15,700. Compared to this incentive strategy, our approach produced an absolute savings of US $10,845 and a relative savings of 69.08%. The second incentive strategy is to provide a flat incentive of US $50 to the 111 eligible respondents who requested compensation which would have cost US $5500. By permitting these 111 respondents to determine their preferred compensation level up to a maximum of US $50, the total value of compensation requested was US $4855. Compared to the second strategy, providing the option to choose the value of compensation to those who requested it resulted in an absolute savings of US $645 and a relative savings of 11.73%.

Determining the appropriate amount of compensation to offer for survey completion remains challenging. Given that the vast majority of respondents who elected compensation took the maximum amount suggests that US $50 may not have been viewed as sufficient by the majority of intensivists requiring a financial incentive to participate in this very short survey. Future studies with the ability to offer greater incentives and, thus, subject to less of a ceiling effect could provide insight into the distribution of preferred compensation for survey participation among physicians.

It is important to consider the ethical and practical ramifications of perceived scarcity and providing different levels of financial compensation to members of the same study cohort. If provided as remuneration for a participant's time, failing to provide sufficient compensation may be ethically untenable or impact data quality. In such cases, a preferable strategy would be to offer an ethically acceptable level of remuneration for all participants completing the survey and to permit the subset of participants who respond most quickly to choose a preferred level of compensation beyond the minimum remuneration level. In such cases, study investigators would need sufficient funds to cover the maximum possible cost of the total compensation and would effectively be incentivizing prompt responses.

Limitations and Strengths

A potential limitation of our study is the generalizability of our results to other groups of physicians and other health care providers. Additionally, as a bundled intervention comprised of three techniques to optimize survey response, we cannot isolate the impact of any one technique or detect any potential synergistic effects. Study strengths include the use of a national database of academic intensivists containing demographic information on physicians who are almost certain to have regular Internet access, and an electronic survey platform that provided

XSL•FO
RenderX

important details regarding survey completion. A study of health care providers offered a gift card to a retail store chain in a postal survey found that provider response decreased in proportion to distance from the nearest store [40]. By offering an incentive that can be used to make Web-based purchases, it is unlikely that the decision to create an incentive was influenced by concerns about proximity to a physical location.

Conclusions

In conclusion, in this randomized controlled trial of 1850 US academic intensivists, giving physicians a time-limited opportunity to choose how much they wished to be compensated for participation in a brief, Web-based survey was associated with a small increase in response rate and a substantial decrease in time to response, without any decrease in how long physicians spent in completing the survey.

Acknowledgments

We wish to thank Cheyenne McCray, BA, and W Chase Hamilton, BA, for their assistance with data collection. This research was supported by the National Institutes of Health (R24 HL111895) and by a Doctoral Thesis Research Fund award from the Department of Epidemiology, Johns Hopkins University Bloomberg School of Public Health. AET received support from the Johns Hopkins University Sommer Scholars Program and a postdoctoral training grant from the National Institute on Aging (T32AG000247).

Authors' Contributions

AET wrote the first draft of the manuscript and is responsible for the overall content as guarantor.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.2 [41].

[PDF File (Adobe PDF File), 223KB - jmir_v17i7e189_app1.pdf]

References

1. Klabunde CN, Willis GB, McLeod CC, Dillman DA, Johnson TP, Greene SM, et al. Improving the quality of surveys of physicians and medical groups: a research agenda. Eval Health Prof 2012 Dec 1;35(4):477-506. [doi: 10.1177/0163278712458283] [Medline: 22947596]
2. Cull WL, O'Connor KG, Sharp S, Tang SS. Response rates and response bias for 50 surveys of pediatricians. Health Serv Res 2005 Feb;40(1):213-226 [FREE Full text] [doi: 10.1111/j.1475-6773.2005.00350.x] [Medline: 15663710]
3. Cook JV, Dickinson HO, Eccles MP. Response rates in postal surveys of healthcare professionals between 1996 and 2005: an observational study. BMC Health Serv Res 2009;9:160 [FREE Full text] [doi: 10.1186/1472-6963-9-160] [Medline: 19751504]
4. Cho YI, Johnson TP, Vangeest JB. Enhancing surveys of health care professionals: a meta-analysis of techniques to improve response. Eval Health Prof 2013 Sep 1;36(3):382-407. [doi: 10.1177/0163278713496425] [Medline: 23975761]
5. Edwards PJ, Roberts I, Clarke MJ, Diguiseppi C, Wentz R, Kwan I, et al. Methods to increase response to postal and electronic questionnaires. Cochrane Database Syst Rev 2009(3):MR000008. [doi: 10.1002/14651858.MR000008.pub4] [Medline: 19588449]
6. Flanigan TS, McFarlane E, Cook S. Conducting survey research among physicians and other medical professionals: a review of current literature. In: Proceedings of the American Association for Public Opinion Research 63rd Annual Conference. 2008 May Presented at: American Association for Public Opinion Research 63rd Annual Conference; May 15-18, 2008; New Orleans, LA p. 4136-4147 URL: http://www.amstat.org/sections/SRMS/proceedings/y2008/Files/flanigan.pdf
7. Asch DA, Jedrziewski MK, Christakis NA. Response rates to mail surveys published in medical journals. J Clin Epidemiol 1997 Oct;50(10):1129-1136. [Medline: 9368521]
8. Halpern SD, Asch DA. Commentary: Improving response rates to mailed surveys: what do we learn from randomized controlled trials? Int J Epidemiol 2003 Aug 1;32(4):637-638 [FREE Full text] [Medline: 12913043]
9. Montori VM, Leung TW, Walter SD, Guyatt GH. Procedures that assess inconsistency in meta-analyses can assess the likelihood of response bias in multiwave surveys. J Clin Epidemiol 2005 Aug;58(8):856-858. [doi: 10.1016/j.jclinepi.2004.11.027] [Medline: 16018920]
10. Listyowardojo TA, Nap RE, Johnson A. Demographic differences between health care workers who did or did not respond to a safety and organizational culture survey. BMC Res Notes 2011;4:328 [FREE Full text] [doi: 10.1186/1756-0500-4-328] [Medline: 21899771]

XSL•FO
RenderX

11. Halbesleben JR, Whitman MV. Evaluating survey quality in health services research: a decision framework for assessing nonresponse bias. Health Serv Res 2013 Jun;48(3):913-930 [FREE Full text] [doi: 10.1111/1475-6773.12002] [Medline: 23046097]
12. McMahon SR, Iwamoto M, Massoudi MS, Yusuf HR, Stevenson JM, David F, et al. Comparison of e-mail, fax, and postal surveys of pediatricians. Pediatrics 2003 Apr;111(4 Pt 1):e299-e303. [Medline: 12671142]
13. Leece P, Bhandari M, Sprague S, Swiontkowski MF, Schemitsch EH, Tornetta P, et al. Internet versus mailed questionnaires: a controlled comparison (2). J Med Internet Res 2004 Oct 29;6(4):e39 [FREE Full text] [doi: 10.2196/jmir.6.4.e39] [Medline: 15631963]
14. Seguin R, Godwin M, MacDonald S, McCall M. E-mail or snail mail? Randomized controlled trial on which works better for surveys. Can Fam Physician 2004 Mar;50:414-419 [FREE Full text] [Medline: 15318679]
15. Nicholls K, Chapman K, Shaw T, Perkins A, Sullivan MM, Crutchfield S, et al. Enhancing response rates in physician surveys: the limited utility of electronic options. Health Serv Res 2011 Oct;46(5):1675-1682 [FREE Full text] [doi: 10.1111/j.1475-6773.2011.01261.x] [Medline: 21492157]
16. Asch DA, Christakis NA, Ubel PA. Conducting physician mail surveys on a limited budget. A randomized trial comparing $2 bill versus $5 bill incentives. Med Care 1998 Jan;36(1):95-99. [Medline: 9431335]
17. Halpern SD, Ubel PA, Berlin JA, Asch DA. Randomized trial of 5 dollars versus 10 dollars monetary incentives, envelope size, and candy to increase physician response rates to mailed questionnaires. Med Care 2002 Sep;40(9):834-839. [doi: 10.1097/01.MLR.0000024722.29926.3D] [Medline: 12218773]
18. Keating NL, Zaslavsky AM, Goldstein J, West DW, Ayanian JZ. Randomized trial of $20 versus $50 incentives to increase physician survey response rates. Med Care 2008 Aug;46(8):878-881. [doi: 10.1097/MLR.0b013e318178eb1d] [Medline: 18665068]
19. Edwards P, Roberts I, Clarke M, DiGuiseppi C, Pratap S, Wentz R, et al. Increasing response rates to postal questionnaires: systematic review. BMJ 2002 May 18;324(7347):1183 [FREE Full text] [Medline: 12016181]
20. Leung GM, Johnston JM, Saing H, Tin KY, Wong IO, Ho LM. Prepayment was superior to postpayment cash incentives in a randomized postal survey among physicians. J Clin Epidemiol 2004 Aug;57(8):777-784. [doi: 10.1016/j.jclinepi.2003.12.021] [Medline: 15485729]
21. Halpern SD, Hussen SA, Metkus TS, Ward NS, Luce JM, Curtis JR. Development of an e-mail database of US intensive care physicians. J Crit Care 2007 Mar;22(1):28-31. [doi: 10.1016/j.jcrc.2006.12.006] [Medline: 17371741]
22. Turnbull AE, Krall JR, Ruhl AP, Curtis JR, Halpern SD, Lau BM, et al. A scenario-based, randomized trial of patient values and functional prognosis on intensivist intent to discuss withdrawing life support. Crit Care Med 2014 Feb 26;42(6):1455-1462. [doi: 10.1097/CCM.0000000000000227] [Medline: 24584065]
23. Qualtrics. URL: http://www.qualtrics.com/ [accessed 2015-07-21] [WebCite Cache ID 6aC4nZsW7]
24. Imai K, King G, Stuart E. Misunderstandings between experimentalists and observationalists about causal inference. J R Stat Soc Ser A Stat Soc 2008;171(2):481-502.
25. US Census Bureau. URL: https://www.census.gov/geo/maps-data/maps/pdfs/reference/us_regdiv.pdf [accessed 2015-07-21] [WebCite Cache ID 6SuG1xhzV]
26. Amazon.com. URL: http://www.amazon.com/ [accessed 2015-07-20] [WebCite Cache ID 6aAXJM9zD]
27. The American Association for Public Opinion Research. Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 7th edition. Deerfield, IL: AAPOR; 2011. URL: http://www.aapor.org/AAPORKentico/AAPOR_Main/media/MainSiteFiles/StandardDefinitions2011_1.pdf [accessed 2015-07-21] [WebCite Cache ID 6aCQ6hO5E]
28. The R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2013 May 16. URL: http://web.mit.edu/r_v3.0.1/fullrefman.pdf [accessed 2015-07-23] [WebCite Cache ID 6aFA9R2wi]
29. Hollander M, Wolfe DA. Nonparametric Statistical Methods. 2nd edition. New York, NY: John Wiley & Sons, Inc; Jan 25, 1999.
30. Lynn M. Scarcity effects on value: a quantitative review of the commodity theory literature. Psychology & Marketing 1991;8(1):43-57. [doi: 10.1002/mar.4220080105]
31. Kahneman D, Knetsch J, Thaler R. Anomalies: the endowment effect, loss aversion, and status quo bias. J Econ Perspect 1991 Feb;5(1):193-206. [doi: 10.1257/jep.5.1.193]
32. Voigt LF, Koepsell TD, Daling JR. Characteristics of telephone survey respondents according to willingness to participate. Am J Epidemiol 2003 Jan 1;157(1):66-73 [FREE Full text] [Medline: 12505892]
33. Chretien J, Chu LK, Smith TC, Smith B, Ryan MA, Millennium Cohort Study Team. Demographic and occupational predictors of early response to a mailed invitation to enroll in a longitudinal health study. BMC Med Res Methodol 2007;7:6 [FREE Full text] [doi: 10.1186/1471-2288-7-6] [Medline: 17397558]
34. Prabhakaran J, Spera C, Leach L, Foster R. Differences in early and late responders: findings from a military Web-based community survey. In: Proceedings of the American Association for Public Opinion Research (AAPOR) 65th Annual Conference. 2010 Presented at: The American Association for Public Opinion Research (AAPOR) 65th Annual Conference; May 11-13, 2010; Chicago, IL URL: http://www.websm.org/db/12/14339/Bibliography/Differences_in_Early_and_Late_Responders_Findings_from_a_Military_WebBased_Community_Survey/

XSL•FO
RenderX

35. Leadbetter S, Hawkins NA, Scholl LE, McCarty FA, Rodriguez JL, Freedner-Maguire N, et al. Recruiting women for a study on perceived risk of cancer: influence of survey topic salience and early versus late response. Prev Chronic Dis 2013;10:E75 [FREE Full text] [doi: 10.5888/pcd10.120293] [Medline: 23660117]
36. Glidewell L, Thomas R, MacLennan G, Bonetti D, Johnston M, Eccles MP, et al. Do incentives, reminders or reduced burden improve healthcare professional response rates in postal questionnaires? two randomised controlled trials. BMC Health Serv Res 2012;12:250 [FREE Full text] [doi: 10.1186/1472-6963-12-250] [Medline: 22891875]
37. Bolt E, van der Heide A, Onwuteaka-Philipsen B. Reducing questionnaire length did not improve physician response rate: a randomized trial. J Clin Epidemiol 2014 Apr;67(4):477-481. [doi: 10.1016/j.jclinepi.2013.09.012] [Medline: 24290148]
38. Cottrell E, Roddy E, Rathod T, Thomas E, Porcheret M, Foster NE. Maximising response from GPs to questionnaire surveys: do length or incentives make a difference? BMC Med Res Methodol 2015;15:3 [FREE Full text] [doi: 10.1186/1471-2288-15-3] [Medline: 25563390]
39. Ziegenfuss JY, Niederhauser BD, Kallmes D, Beebe TJ. An assessment of incentive versus survey length trade-offs in a Web survey of radiologists. J Med Internet Res 2013 Mar 20;15(3):e49 [FREE Full text] [doi: 10.2196/jmir.2322] [Medline: 23514869]
40. Van Otterloo J, Richards JL, Seib K, Weiss P, Omer SB. Gift card incentives and non-response bias in a survey of vaccine providers: the role of geographic and demographic factors. PLoS One 2011;6(11):e28108 [FREE Full text] [doi: 10.1371/journal.pone.0028108] [Medline: 22132224]
41. Eysenbach G, Consort- E. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

eC: the estimated proportion in the control arm
eI: the estimated proportion in the intervention arm
I: number of eligible participants who responded to all survey questions
ICU: intensive care unit
IQR: interquartile range
IRB: Institutional Review Board
N/A: not applicable
P: eligible participants who answered the eligibility screening question but did not answer all survey questions
RR3: response rate three
UH: nonresponders
UO: responders with unknown eligibility

Edited by G Eysenbach; submitted 28.09.14; peer-reviewed by S Leadbetter, A Naidech; comments to author 05.03.15; revised version received 27.03.15; accepted 03.07.15; published 29.07.15.

Please cite as:
Turnbull AE, O'Connor CL, Lau B, Halpern SD, Needham DM
Allowing Physicians to Choose the Value of Compensation for Participation in a Web-Based Survey: Randomized Controlled Trial
J Med Internet Res 2015;17(7):e189
URL: http://www.jmir.org/2015/7/e189/
doi:10.2196/jmir.3898
PMID:26223821

©Alison E Turnbull, Cristi L O'Connor, Bryan Lau, Scott D Halpern, Dale M Needham. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 29.07.2015. 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

Information and Communication Technology Use Among Low-Income Pregnant and Postpartum Women by Race and Ethnicity: A Cross-Sectional Study

Nymisha Chilukuri[1], BSc; Meredith West[1], MSPH; Janice Lynn Henderson[2], MD; Shari Lawson[2], MD; Robert Ehsanipoor[2], MD; Kathleen Costigan[2], RN; Sarah Polk[3], ScM, MD; Wendy Bennett[1], MD, MPH

[1]Division of General Internal Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, United States

[2]Division of General Obstetrics and Gynecology, Johns Hopkins University School of Medicine, Baltimore, MD, United States

[3]Division of General Pediatrics, Johns Hopkins University School of Medicine, Baltimore, MD, United States

Corresponding Author:
Wendy Bennett, MD, MPH
Division of General Internal Medicine
Johns Hopkins University School of Medicine
Suite 2-616
2024 E. Monument St
Baltimore, MD, 21205
United States
Phone: 1 410 502 6081
Fax: 1 410 955 0476
Email: wendy.bennett@jhmi.edu

Abstract

Background: Pregnancy and the postpartum period provide windows of opportunity to impact perinatal and lifelong preventive health behavior for women and their families, but these opportunities are often missed. Understanding racial/ethnic differences in information and communication technology (ICT) use could inform technology-based interventions in diverse populations.

Objective: The objective of the study was to evaluate differences in the use of ICT between racial and ethnic groups as well as by English language proficiency.

Methods: We conducted a cross-sectional study of 246 women who were aged 18 years or older and pregnant or within 1 year of delivery. They were recruited from 4 hospital-based outpatient clinics and completed a self-administered survey. We used multivariate regression analysis to evaluate the association between race/ethnicity and ICT (mobile phone/short message service [SMS] text message, Internet, and social network) usage by race/ethnicity and perceived English language proficiency after adjusting for age, income, marital status, and insurance status.

Results: In all, 28% (69/246) of participants were Latina, 40% (98/246) were African American, 23% (56/246) were white, and 9% (23/246) from other racial/ethnic groups. Of the Latinas, 84% (58/69) reported limited English language proficiency and 59% (41/69) were uninsured. More than 90% of all participants reported mobile phone use, but more than 25% (65/246) had changed phone numbers 2 or more times in the past year. Compared to white women, African American women were less likely to SMS text message (OR 0.07, 95% CI 0.01-0.63) and Latinas were less likely to use the Internet to find others with similar concerns (OR 0.23, 95% CI 0.08-0.73). Women with limited English language proficiency were less likely to use the Internet overall (OR 0.30, 95% CI 0.09-0.99) or use email (OR 0.22, 95% CI 0.08-0.63) compared to women with adequate English language proficiency.

Conclusions: Mobile phones are widely available for the delivery of health interventions to low-income, racially diverse pregnant and postpartum women, but disparities in Internet use and SMS text messaging exist. Interventions or programs requiring Web-based apps may have lower uptake unless alternatives are available, such as those adapted for limited English proficiency populations.

(J Med Internet Res 2015;17(7):e163) doi:10.2196/jmir.3916

XSL•FO
RenderX

KEYWORDS

pregnancy; health services accessibility; postpartum period; cell phones; text messaging; Internet; health status disparities; Hispanic Americans

Introduction

Information and communication technologies (ICT), including mobile phones (eg, smartphones and regular mobile phones), Internet, email, and social networking have the potential to improve care for underserved communities with reduced access to health care. Although pregnant women frequently utilize health care services, they are often lost to follow-up after delivery, even among those with high-risk pregnancies [1]. Prior studies in pregnant and postpartum women show increasing use of Internet for delivery of health information [2] and interest in mobile phone apps [3]. Potential participants in postpartum weight loss interventions report high interest in Web-based components [4]. ICT has the potential to engage with and educate women, thus promoting improved health for women and their families before and after delivery [5], but evidence to support the uptake among low-income racially and ethnically diverse populations is limited.

The Pew Research Center showed 90% of Americans used a mobile phone [6] and 87% reported at least some Internet use [7]. Among those who report Internet use, 91% send or read emails and 74% use social networking sites such as Facebook, Myspace, or LinkedIn. Those with higher education and in younger age groups report most frequent Internet use, suggesting that ICT may be an optimal tool for communication, monitoring, education, and even providing interventions to young women [7]. However, studies have shown that low-income populations have fewer available ICT resources, including a reliable home Internet connection, and also frequently change mobile phone numbers [8,9]. People with limited English proficiency face multiple barriers to health care and community services, indicating a high need for novel and effective outreach strategies to engage and communicate with this population [10-14].

To inform the development of ICT-based health-related programming to a diverse population of pregnant and postpartum women, we conducted a cross-sectional survey about their ICT usage. The aim of this study was to evaluate differences in the use of ICT between racial and ethnic groups and by English language proficiency. We hypothesized that pregnant and postpartum women frequently use ICT, including mobile phones, Internet, and social networking, and that rates of ICT use would be lowest among Latinas and those with limited English language proficiency. The rationale for the hypothesis was based on prior literature indicating that Latino adults, especially those with less education, had lower Internet usage [15].

We designed the study to inform the development of culturally appropriate ICT interventions to promote healthy lifestyle behaviors in the perinatal period.

Methods

Study Design

This was a cross-sectional study, using a one-time self-administered questionnaire, to describe ICT use among women who were pregnant or in the first year postpartum. The study was approved by the Johns Hopkins University institutional review board.

Study Setting and Population

We recruited 246 women who were attending a clinical visit at 1 of 4 outpatient obstetric or pediatric clinical sites from 2 hospitals in Baltimore, MD, between January and April 2013. Three of the sites provided high-risk obstetric care. Women were eligible to participate if they were aged at least 18 years, reported that they could read English or Spanish either “well” or “very well” and were either pregnant or within 1 year of delivery.

Survey Design and Data Collection

We designed a 68-item questionnaire with items adapted from validated instruments to assess sociodemographics, use of ICT, and self-efficacy (confidence on their ability) for accessing online health information. Self-efficacy was assessed using the question “How confident are you in your ability to find helpful and useful health information on the Internet?” adapted from the Perceived Efficacy in Patient-Physician Interactions (PEPPI) 5-item scale [16]. The stem of these items began with “How confident are you in your ability to...” which we adapted to be specific to health information on the Internet. The questions used a 5-point response scale (1=“not at all confident” and 5=“extremely confident”), which we dichotomized based on the distribution of responses into “extremely confident” and “somewhat confident” versus “neutral,” “not very,” and “not at all confident.” Questions on medical history and access to care were adapted from several national surveys including the Center for Disease Control and Prevention’s Behavioral Risk Factor Surveillance System [11] or the Pregnancy Risk Assessment Monitoring System Core and Standard Questionnaires [12]. ICT usage questions were adapted from The Pew Research Center’s questionnaires on Peer-to-Peer Healthcare [13] and Health Online [14]. The final survey was translated into Spanish. The English and Spanish versions of the survey were pilot-tested among English- and Spanish-speaking patients to ensure cultural relevance, understandability, readability (aiming for fifth grade reading level or less), and completion within 10-15 minutes. Eligible participants completed a 10-15 minute self-administered questionnaire either immediately before or after their outpatient clinic visit in a private space. We offered an audio-recorded version that read each question aloud using a CD player, but no one chose this version. Participants received a US $10 gift card.

Measures

The primary outcome was use of ICT, which included use of mobile phone, short message service (SMS) text messaging,

XSL•FO
RenderX

Internet, email, and social networks. We also asked whether participants used these technologies to identify health information for themselves and their families and, if so, what they found useful.

The main independent variables were self-reported race/ethnicity, coded as non-Hispanic black, non-Hispanic white, Hispanic (Latino), and other races, and self-perceived spoken English language proficiency. English language proficiency was assessed based on the US Census question and categorization, which had been incorporated into the Consumer Assessment of Healthcare Providers and Systems (CAHPS) Cultural Competence Supplemental Survey [17,18]: "How well do you speak English?" The response of "very well" was coded as "adequate" and any response of less than "very well" ("well," "not well," and "not at all") was coded as having "limited English proficiency" based on previous validation of these cutpoints [17-19].

Other descriptors and covariates included sociodemographic variables [17] (eg, education level, income, marital status), pregnancy status, self-reported medical history, and health insurance.

Data Analysis

We used descriptive statistics (*t* tests for continuous variables and chi-square tests for categorical variables) to assess and describe the characteristics of our sample and use of ICT by race/ethnicity. We created multivariate logistic regression models to assess the association between race and ethnicity and mode of ICT after adjusting for age and education level.

Results

Characteristics of Study Sample

The mean age in our sample of 246 women was 28 (SD 6) years with white women being slightly older (mean 31, SD 6 years) than Latina (mean 28, SD 6 years) and African American (mean 26, SD 6 years) women. Most women were pregnant at the time of the survey.In all, 28% (69/246) were Latina, 40% (98/246) were African American, 23% (56/246) were white, and 9% (23/246) were from other racial/ethnic groups, which included Asian (n=10), Native Hawaiian and Pacific Islander (n=1), American Indian/Alaskan Native (n=4), and multiethnic women (n=8). In all, 17% (12/69) of Latinas and 4% (33/98) of African Americans reported household incomes less than US $10,000 compared to 9% (5/56) of white women. For insurance status, 54% (132/246) were insured with Medicaid or Medicare, but 60% (41/69) of Latinas were uninsured and 36% (89/246) of women were employed either full or part time. For Latina women, 84% (58/69) reported limited spoken English language proficiency compared to 1.1% (2/177) of the other racial/ethnic groups. Latinas most commonly reported Mexico (29%, 20/69) and El Salvador (28%, 19/69) as countries of origin. The sample had a high prevalence of medical conditions including type 2 diabetes (7%, 16/246), gestational diabetes (11%, 28/246), hypertension (12%, 29/246), and overweight/obesity (56%, 138/246). For white and African American women, 84% (47/56) and 72% (71/98), respectively, reported having a primary care physician compared to 19% (13/69) of Latina women (Table 1).

XSL•FO
RenderX

Table 1. Characteristics of the sample of pregnant and postpartum women (N=246).

Characteristics	Total N=246	Latino n=69	African American n=98	White n=56	Other races n=23	*P*
Age (years), mean (SD)	28 (6)	28 (6)	26 (6)	31 (6)	29 (7)	<.001
Married or living with partner, n (%)	173 (70)	59 (86)	50 (51)	44 (79)	20 (87)	<.001
Currently pregnant, n (%)	206 (84)	53 (77)	85 (87)	50 (89)	18 (78)	.18
Limited English proficiency[a]	60 (24)	58 (84)	1 (1)	1 (2)	0 (0)	<.001
Household income (US $), [c] n (%)						<.001
<$10,000	53 (22)	12 (17)	33 (34)	5 (9)	3 (13)	
$10,000-$49,999	83 (34)	26 (38)	37 (38)	14 (25)	6 (26)	
>$50,000	50 (20)	4 (6)	4 (4)	30 (54)	12 (52)	
Education, n (%)						<.001
<Grade 12 or GED	60 (24)	36 (52)	13 (13)	7 (13)	4 (17)	
Grade 12 or GED	83 (34)	22 (32)	48 (49)	10 (18)	3 (13)	
>Grade 12 or GED	101 (41)	10 (14)	36 (37)	39 (70)	16 (70)	
Insurance status, n (%)						<.001
Commercial plan	66 (27)	7 (10)	14 (14)	32 (57)	13 (57)	
Medicaid/Medicare[c]	132 (54)	19 (28)	80 (82)	24 (43)	9 (39)	
Uninsured	45 (18)	41 (59)	3 (3)	0 (0)	1 (4)	
Employment status, n (%)						.002
Employed	89 (36)	15 (22)	35 (36)	27 (48)	12 (52)	
Homemaker/maternity leave	68 (28)	33 (48)	16 (16)	14 (25)	5 (22)	
Attending school	13 (5)	2 (3)	6 (6)	3 (5)	2 (9)	
Unemployed	69 (28)	15 (22)	38 (39)	2 (4)	4 (17)	
Medical history, [d] n (%)						
Type 2 diabetes	16 (7)	3 (4)	10 (10)	2 (4)	1 (4)	.30
Gestational diabetes	28 (11)	7 (10)	8 (8)	12 (21)	1 (4)	.05
High blood pressure	29 (12)	4 (6)	13 (13)	8 (14)	4 (17)	.31
Overweight or obese[e]	138 (56)	30 (44)	70 (71)	27 (48)	11 (48)	.001
Has primary care physician	147 (60)	13 (19)	71 (72)	47 (84)	16 (70)	<.001
Phone and Internet use						
Uses mobile phone	234 (95)	65 (94)	90 (92)	56 (100)	23 (100)	.15
Uses smartphone	172 (74)	38 (55)	69 (77)	47 (84)	18 (78)	.004
≥2 mobile phone numbers in last 12 months	65 (26)	17 (25)	34 (35)	10 (18)	4 (17)	.08
Has home phone	106 (43)	25 (36)	48 (49)	24 (43)	9 (39)	.69
Uses Internet	209 (85)	43 (62)	90 (92)	54 (96)	22 (96)	<.001
High (vs low) self-efficacy for using Internet	145 (59)	20 (29)	62 (63)	44 (79)	19 (83)	<.001
ICT outcomes						
SMS text messaging	222 (90)	61 (88)	83 (85)	55 (98)	23 (100)	.02
Email	193 (79)	35 (51)	84 (86)	52 (93)	22 (96)	<.001
Internet	209 (85)	43 (62)	90 (92)	54 (96)	22 (96)	<.001
Use of Internet to find health info	182 (74)	35 (51)	77 (79)	49 (88)	21 (91)	<.001

XSL•FO
RenderX

Characteristics	Total N=246	Latino n=69	African American n=98	White n=56	Other races n=23	*P*
Using Internet to find others with similar concerns	102 (42)	8 (12)	48 (49)	36 (64)	10 (43)	<.001
Social networking	187 (76)	37 (54)	81 (83)	49 (88)	20 (87)	<.001

[a] Limited language proficiency defined less than "very well" on the question "How well do you speak English?"

[b] Survey item provided the option of declining to disclose income. A total of 24% (60/246) declined: 39% (27/69) Latino, 25% (24/98) African American, 13% (7/56) white, 9% (7/23) other races.

[c] Proportion with Medicaid (vs Medicare) was 52% (128/246) with 26% (18/69) Latinas, 77% (77/98) African American, 43% (24/56) white, and 39% (9/23) with other races.

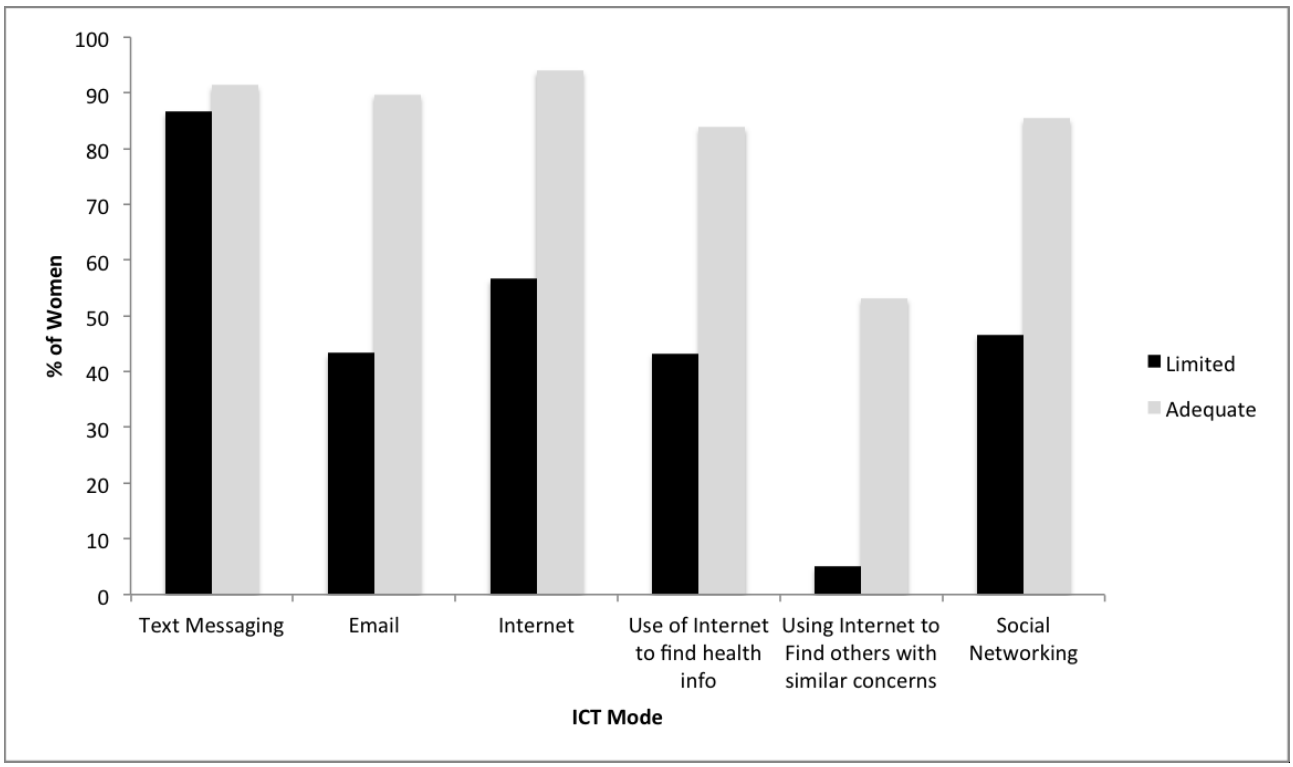
[d] Medical history is self-reported.

[e] Body mass index (BMI) calculated using self-reported prepregnancy weight (kg) divided by the square of self-reported height (m^2). Overweight or obese defined as BMI ≥30 kg/m^2.

Rates of Information and Communication Technology Usage

Mobile phone use was greater than 90% (234/246) among all racial and ethnic groups with African American women reporting the lowest rate of 92% (90/98) (Table 1). Compared with African American and white women, fewer Latina women used smartphones (55%, 38/69), social networking sites (54%, 37/69), or accessed the Internet (62%, 43/69) (Table 1). However, the majority of women in all racial/ethnic groups used mobile phones for SMS text messaging, although the rate was slightly lower for African American women (Latina: 88%, 61/69; African American: 85%, 83/98; white: 98%, 55/56) (Table 1). More than one-quarter of the sample (26%, 65/246) reported having 2 or more different mobile phone numbers in the past 12 months and 43% (106/246) of women reported having a home phone number or landline. Compared with women with adequate spoken English language proficiency, women with limited English language proficiency less frequently used all forms of ICT (Figure 1). Among Latinas, those with limited English proficiency had lower use of Internet (38/69, 55%) compared to Latinas with adequate English language proficiency (55%, 32/58 vs 100%, 11/11, *P*=.005) (not shown).

Figure 1. Rates of use of information and communication technology modality and function by English language proficiency (adequate vs limited).



XSL•FO
RenderX

Odds of Information and Communication Technology Use by Race/Ethnicity and English Language Proficiency

Compared to a white reference group, African American women were statistically significantly less likely to report SMS text messaging (OR 0.08, 95% CI 0.01-0.67) after adjustment for age and education (Table 2). Compared to white women, Latinas were less likely to report using the Internet (OR 0.15, 95% CI 0.03-0.78), email (OR 0.17, 95% CI 0.05-0.61), social networking (OR 0.27, 95% CI 0.09-0.75), Internet use to find other people with similar health concerns (OR 0.16, 95% CI 0.06-0.43), and had lower self-efficacy for health information–related Internet use (OR 0.26, 95% CI 0.10-0.68) (Table 2).

Table 2. Odds of using information and communication technology (ICT) by race/ethnicity.[a]

ICT use	Latino, OR (95% CI) n=69	African American, OR (95% CI) n=98
SMS text message	0.18 (0.02-1.61)	0.08 (0.01-0.67)
Internet use[b]	0.15 (0.03-0.78)	0.54 (0.10-3.04)
Email use	0.17 (0.05-0.61)	0.52 (0.14-1.91)
Social networking	0.27 (0.09-0.75)	0.58 (0.20-1.66)
Internet used to obtain health Information[b]	0.38 (0.14-1.07)	0.90 (0.32-2.54)
Internet used to find others with similar concerns	0.16 (0.06-0.43)	0.86 (0.39-1.89)
High (vs low) self-efficacy for Internet use[c]	0.26 (0.10-0.68)	0.62 (0.26-1.49)

[a] Reference=white race. Model adjusted for age and education.

[b] Includes accessing Internet via mobile phone or computer.

[c] Self-efficacy assessed using question “How confident are you in your ability to find helpful and useful health information on the Internet?” and categorized as high=extremely confident and somewhat confident vs low=neutral, not very, and not at all confident.

Compared to women with adequate English language proficiency, women with lower English language proficiency were equally likely to SMS text message (OR 0.97, 95% CI 0.34-2.72), but had a lower likelihood of using the Internet (OR 0.20, 95% CI 0.08-0.47), email (OR 0.19, 95% CI 0.09-0.41), social networking (OR 0.27, 95% CI 0.13-0.57), Internet used to obtain health information (OR 0.27, 95% CI 0.13-0.56), and Internet used to find others with similar concerns (OR 0.08, 95% CI 0.02-0.28) (Table 3).

Table 3. Odds of information and communication technology (ICT) use for women with low vs adequate English proficiency.[a]

ICT use	Low English language proficiency, OR (95% CI)
SMS text message	0.97 (0.34-2.72)
Internet use[b]	0.20 (0.08-0.47)
Email use	0.19 (0.09-0.41)
Social networking	0.27 (0.13-0.57)
Internet used to obtain health information[b]	0.27 (0.13-0.56)
Internet used to find others with similar concerns	0.08 (0.02-0.28)
High (vs low) self-efficacy for Internet use[c]	0.21 (0.09-0.45)

[a] Model adjusted for age and education.

[b] Includes accessing Internet via mobile phone or computer.

[c] Self-efficacy assessed using question “How confident are you in your ability to find helpful and useful health information on the Internet?” and categorized as high=extremely confident and somewhat confident vs low=neutral, not very, and not at all confident.

Discussion

In this sample of low-income, racially and ethnically diverse pregnant and postpartum women, mobile phone and SMS text message usage were common across all racial/ethnic groups. Although more than 85% of all participants reported SMS text messaging, African American women were less likely to text compared to white women. The rates of mobile phone usage reported in our survey were similar to The Pew Research Center’s Internet Project Survey data showing 90% of people own cell phones, including 90% of whites and African Americans and 92% of Hispanics [6]. Additionally, we confirmed low rates of landline use (57% did not own landlines) consistent with Pew Internet Research findings (41%). In the Pew survey, households with lower socioeconomic status and

XSL•FO
RenderX

Hispanics had higher rates of not having landlines (56.2% and 53.1%, respectively) [20]. Although our sample generally reported high mobile phone and SMS text messaging rates, we identified disparities in Internet, email, and social networking use by racial/ethnic groups and limited English language proficiency. Our results suggest that mobile phones are potentially useful modalities for the delivery of health interventions to low-income pregnant and postpartum women, but interventions requiring Web-based apps may have lower uptake unless alternatives (ie, paper) and Spanish translations are available.

Other studies have also reported lower rates of Internet usage among Latinas, but without close examination of the role of English language proficiency. A large cross-sectional survey of 3181 young women attending reproductive health clinics in Texas reported 92.7% of whites and 92.9% of African American women, but only 67.5% of Latinos, used the Internet. Hispanic women reported barriers to Internet use including cost, not having a computer at home, and not knowing how to use a computer [21]. However, the impact of English language proficiency on differential ICT use was not described. Notably, more than 80% of Latinas in our study had limited English language proficiency and 17% spoke any English, which is in contrast to the Latino population sampled in the Pew Internet Research Survey in which 65% were English speaking (either English dominant or bilingual) and proficiency was not assessed. The large difference in English proficiency likely accounts for the disparity we noted in Internet use, in which only 62% of Latinas in our sample used the Internet compared to 78% of Latinos in the Pew Survey [22]. In addition, more than half of the Latinas in our sample reported less than a high school education; from other surveys from this community, more than one-third likely have less than a sixth grade education [15] indicating lower literacy, including Spanish literacy. Our results suggest that English language and literacy are major barriers for women to use and access the Internet. The development of technology-based interventions, especially those that require Internet components, should be translated into Spanish, designed for people with lower literacy, and culturally adapted for Latinos to have the greatest potential impact.

Despite these disparities in Internet usage, our study supports mobile phone-based interventions in a low-income, racially diverse population. Growing evidence supports mobile phone-based interventions to impact health behaviors, but few studies have focused on pregnancy and postpartum health [23]. One example of a large-scale SMS text messaging program aimed at improving prenatal care and pregnancy outcomes is the Text4Baby program launched in 2010 by the Centers for Disease Control and Prevention. Women who sign up receive texts containing information about prenatal and postpartum health behaviors and services [24]. However, a randomized controlled trial of 123 women (approximately 80% of whom were Spanish speaking and 75.6% who had participated in the Special Supplemental Nutrition Program for Women, Infants, and Children [WIC] program) did not show a difference between text4baby intervention and usual care control in terms of changes in self-reported health behaviors, but no birth or utilization of care outcomes were reported [25]. Adaptation of the SMS text messaging programs for low-income women is especially important. The WIC program serves low-income women and children and has also been focused on improving care delivery through mobile-based apps [25,26]. A WIC program in Atlanta, GA, tested the text4baby program to assess participants' enrollment and satisfaction in 468 (91% African American) participants. Only 51% of women provided with enrollment instructions attempted to enroll in the program; among these, 69% successfully enrolled mostly via SMS text message (vs online). Higher education and higher incomes were associated with increased enrollment, indicating that the enrollment process may have more barriers for less educated and poorer women [27]. This study notes the importance of testing actual use of ICT interventions in low-income populations to reduce their risk of widening the disparities that they were designed to address.

An additional challenge to implementing and sustaining SMS text messaging interventions and programs is the frequency with which women reported changing mobile phones or phone numbers. In our sample, more than one-quarter of women reported having 2 or more cell phone numbers in the last 12 months, with the highest rates among African American women (35%) and Latinas (25%). To facilitate intervention adherence involving use of mobile phones, studies have budgeted funding to provide mobile phone minutes or plans or even phones to participate to enhance participation rates, but this may not be cost-effective for community-based programs and the aforementioned intervention was not offered in multiple languages [28]. Other studies have required an unlimited short messaging plan [29], but this may exclude lower income populations [30].

The major strengths of our study are including both obstetric and pediatric sites, and identifying a higher risk population of women with medical comorbidities and a racially and ethnically diverse sample. We also collected information about perceived English language proficiency to analyze the results not only by racial and ethnic groups, but also by English language proficiency.

There are several limitations of our study. First, we surveyed a convenience sample of women who were attending one of several clinical sites and may have missed women who did not receive prenatal care or who did not attend visits. This may have made the study's results less generalizable to other women in Baltimore, but because we collected data over several months, women had multiple opportunities to attend visits and complete the surveys. Second, because this was a self-administered questionnaire, we screened out 5 participants who had self-reported low literacy in English or Spanish as part of study eligibility and thus our results provide a "best-case scenario" for Internet and mobile phone use among literate women. Third, the cross-sectional design limits our ability to assess causality. Fourth, our study examined racial and ethnic differences in Internet and mobile phone usage in a population of low-income women in Baltimore, MD, and these differences may be different in other cities in the United States. Fifth, there is a potential that results being attributed to ethnicity may in fact be due to income instead, especially because the latter was not included in the multivariate regression analysis. Sixth, our study did not account

XSL•FO
RenderX

for potential differences in ICT usage for nulliparous versus multiparous women because more experienced mothers may differ in their level of need for information.

In conclusion, our findings show that the racial and ethnic digital divide regarding mobile phone use and SMS text messaging is diminishing, but persists for Internet, email, and social networking by race and ethnicity and particularly for women with limited English language proficiency. These findings support developing linguistically and culturally appropriate mobile phone and SMS text messaging interventions for women of all ethnicities and language proficiencies to promote improved healthy lifestyle behaviors, specifically in the perinatal period.

Acknowledgments

Dr Wendy Bennett is supported by a career development award from the National Heart, Lung, and Blood Institute, 5K23HL098476-02.

Conflicts of Interest

None declared.

References

1. Bennett WL, Chang HY, Levine DM, Wang L, Neale D, Werner EF, et al. Utilization of primary and obstetric care after medically complicated pregnancies: an analysis of medical claims data. J Gen Intern Med 2014 Apr;29(4):636-645 [FREE Full text] [doi: 10.1007/s11606-013-2744-2] [Medline: 24474651]
2. Nicklas JM, Zera CA, England LJ, Rosner BA, Horton E, Levkoff SE, et al. A web-based lifestyle intervention for women with recent gestational diabetes mellitus: a randomized controlled trial. Obstet Gynecol 2014 Sep;124(3):563-570. [doi: 10.1097/AOG.0000000000000420] [Medline: 25162257]
3. Hearn L, Miller M, Fletcher A. Online healthy lifestyle support in the perinatal period: what do women want and do they use it? Aust J Prim Health 2013;19(4):313-318. [doi: 10.1071/PY13039] [Medline: 23899373]
4. Nicklas JM, Zera CA, Seely EW, Abdul-Rahim ZS, Rudloff ND, Levkoff SE. Identifying postpartum intervention approaches to prevent type 2 diabetes in women with a history of gestational diabetes. BMC Pregnancy Childbirth 2011;11:23 [FREE Full text] [doi: 10.1186/1471-2393-11-23] [Medline: 21435246]
5. Jordan ET, Ray EM, Johnson P, Evans WD. Text4Baby: using text messaging to improve maternal and newborn health. Nurs Womens Health 2011;15(3):206-212. [doi: 10.1111/j.1751-486X.2011.01635.x] [Medline: 21672170]
6. Pew Research Center. 2014. Mobile technology fact sheet URL: http://www.pewinternet.org/fact-sheets/mobile-technology-fact-sheet/ [accessed 2014-10-03] [WebCite Cache ID 6T2Myp3Sg]
7. Pew Research Center. 2014. Internet user demographics URL: http://www.pewinternet.org/data-trend/internet-use/latest-stats/ [accessed 2014-10-03] [WebCite Cache ID 6T2N9QLIK]
8. Samal L, Hutton HE, Erbelding EJ, Brandon ES, Finkelstein J, Chander G. Digital divide: variation in internet and cellular phone use among women attending an urban sexually transmitted infections clinic. J Urban Health 2010 Jan;87(1):122-128 [FREE Full text] [doi: 10.1007/s11524-009-9415-y] [Medline: 19941085]
9. Mayben JK, Giordano TP. Internet use among low-income persons recently diagnosed with HIV infection. AIDS Care 2007 Oct;19(9):1182-1187. [doi: 10.1080/09540120701402806] [Medline: 18058404]
10. Centers for Disease Control and Prevention. 2009. Infant feeding practices study II and its year six follow-up URL: http://www.cdc.gov/ifps/ [accessed 2014-10-03] [WebCite Cache ID 6T2PgJovD]
11. Centers for Disease Control and Prevention. 2013. Behavioral risk factor surveillance system URL: http://www.cdc.gov/brfss/ [accessed 2014-10-03] [WebCite Cache ID 6T2PYvg5Q]
12. Center for Disease Control and Prevention. 2009. Pregnancy risk assessment monitoring system URL: http://www.cdc.gov/prams/ [accessed 2014-10-03] [WebCite Cache ID 6T2Pn5dho]
13. Fox S. Peer-to-Peer Healthcare. Washington, DC: Pew Internet & American Life Project; 2011 Feb 28. URL: http://www.pewinternet.org/2011/02/28/peer-to-peer-health-care-2/ [accessed 2014-10-03] [WebCite Cache ID 6T2Q2Hc2R]
14. Fox S, Duggan M. Health Online 2013. Washington, DC: Pew Internet & American Life Project; 2013 Jan 15. URL: http://www.pewinternet.org/2013/01/15/health-online-2013/ [accessed 2014-10-02] [WebCite Cache ID 6T2Q7cN5t]
15. Leite L, Buresh M, Rios N, Conley A, Flys T, Page KR. Cell phone utilization among foreign-born Latinos: a promising tool for dissemination of health and HIV information. J Immigr Minor Health 2014 Aug;16(4):661-669. [doi: 10.1007/s10903-013-9792-x] [Medline: 23440452]
16. Maly RC, Frank JC, Marshall GN, DiMatteo MR, Reuben DB. Perceived efficacy in patient-physician interactions (PEPPI): validation of an instrument in older persons. J Am Geriatr Soc 1998 Jul;46(7):889-894. [Medline: 9670878]
17. Shin H, Bruno R. Language Use and English-Speaking Ability: 2000. Washington, DC: US Census Bureau; 2003. URL: http://www.census.gov/prod/2003pubs/c2kbr-29.pdf [accessed 2014-10-03] [WebCite Cache ID 6T2NM6lwR]
18. Agency for Healthcare Research and Quality. CAHPS Cultural Competence Item Set URL: https://cahps.ahrq.gov/surveys-guidance/item-sets/cultural/index.html [accessed 2014-10-03] [WebCite Cache ID 6T2QDgBLo]

XSL•FO
RenderX

19. Kominski R. How good is 'how well'? an examination of the Census English-speaking ability question. 1989 Aug 06 Presented at: Annual Meeting of the American Statistical Association; August 6, 1989; Washington, DC URL: https://www.census.gov/population/socdemo/language/ASApaper1989.pdf
20. Desilver D. Pew Research Center. 2014 Jul 08. CDC: Two of every five US households have only wireless phones URL: http://www.pewresearch.org/fact-tank/2014/07/08/two-of-every-five-u-s-households-have-only-wireless-phones/ [accessed 2015-03-26] [WebCite Cache ID 6XJ5k88B6]
21. Laz TH, Berenson AB. Racial and ethnic disparities in internet use for seeking health information among young women. J Health Commun 2013;18(2):250-260 [FREE Full text] [doi: 10.1080/10810730.2012.707292] [Medline: 23130608]
22. Lopez M, Gonzalez-Barrera A, Patten E. Pew Research Center. 2013 Mar 07. Closing the digital divide: Latinos and technology adoption 2013 URL: http://www.pewhispanic.org/2013/03/07/closing-the-digital-divide-latinos-and-technology-adoption/ [accessed 2015-03-26] [WebCite Cache ID 6XJ674V5U]
23. 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 Jan;10(1):e1001362 [FREE Full text] [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]
24. Text4Baby. URL: https://www.text4baby.org/ [accessed 2014-10-03] [WebCite Cache ID 6T2Nebhtp]
25. Evans WD, Wallace JL, Snider J. Pilot evaluation of the text4baby mobile health program. BMC Public Health 2012;12:1031 [FREE Full text] [doi: 10.1186/1471-2458-12-1031] [Medline: 23181985]
26. Kraschnewski JL, Chuang CH, Poole ES, Peyton T, Blubaugh I, Pauli J, et al. Paging "Dr. Google": does technology fill the gap created by the prenatal care visit structure? Qualitative focus group study with pregnant women. J Med Internet Res 2014 Jun;16(6):e147 [FREE Full text] [doi: 10.2196/jmir.3385] [Medline: 24892583]
27. Gazmararian JA, Elon L, Yang B, Graham M, Parker R. Text4baby program: an opportunity to reach underserved pregnant and postpartum women? Matern Child Health J 2014 Jan;18(1):223-232. [doi: 10.1007/s10995-013-1258-1] [Medline: 23494485]
28. Song H, May A, Vaidhyanathan V, Cramer EM, Owais RW, McRoy S. A two-way text-messaging system answering health questions for low-income pregnant women. Patient Educ Couns 2013 Aug;92(2):182-187. [doi: 10.1016/j.pec.2013.04.016] [Medline: 23711635]
29. Abroms LC, Boal AL, Simmens SJ, Mendel JA, Windsor RA. A randomized trial of Text2Quit: a text messaging program for smoking cessation. Am J Prev Med 2014 Sep;47(3):242-250. [doi: 10.1016/j.amepre.2014.04.010] [Medline: 24913220]
30. Arora S, Peters AL, Burner E, Lam CN, Menchine M. Trial to examine text message-based mHealth in emergency department patients with diabetes (TExT-MED): a randomized controlled trial. Ann Emerg Med 2014 Jun;63(6):745-54.e6. [doi: 10.1016/j.annemergmed.2013.10.012] [Medline: 24225332]

Abbreviations

BMI: body mass index
CAHPS: Consumer Assessment of Healthcare Providers and Systems
ICT: information and communication technology
PEPPI: Perceived Efficacy in Patient-Physician Interactions
SMS: short message service
WIC: Special Supplemental Nutrition Program for Women, Infants, and Children

Edited by G Eysenbach; submitted 05.10.14; peer-reviewed by J Nicklas, S McRoy; comments to author 15.01.15; revised version received 26.03.15; accepted 19.04.15; published 03.07.15.

Please cite as:
Chilukuri N, West M, Henderson JL, Lawson S, Ehsanipoor R, Costigan K, Polk S, Bennett W
Information and Communication Technology Use Among Low-Income Pregnant and Postpartum Women by Race and Ethnicity: A Cross-Sectional Study
J Med Internet Res 2015;17(7):e163
URL: http://www.jmir.org/2015/7/e163/
doi:10.2196/jmir.3916
PMID:26142162

©Nymisha Chilukuri, Meredith West, Janice Lynn Henderson, Shari Lawson, Robert Ehsanipoor, Kathleen Costigan, Sarah Polk, Wendy Bennett. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 03.07.2015. 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,

XSL•FO
RenderX

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

Prevalence and Frequency of mHealth and eHealth Use Among US and UK Smokers and Differences by Motivation to Quit

Belinda Borrelli[1,2], PhD; Yvonne Kiera Bartlett[2], PhD; Erin Tooley[3], PhD; Christopher J Armitage[2], PhD; Alison Wearden[2], PhD

[1]Boston University, Henry M Goldman School of Dental Medicine, Boston, MA, United States

[2]Manchester Centre for Health Psychology, School of Psychological Sciences, Manchester Academic Health Science Centre, University of Manchester, Manchester, United Kingdom

[3]Feinstein College of Arts & Sciences, Roger Williams University, Bristol, RI, United States

Corresponding Author:
Belinda Borrelli, PhD
Boston University
Henry M Goldman School of Dental Medicine
560 Harrison Ave, 3rd floor
Boston, MA, 02118
United States
Phone: 1 617 414 1116
Fax: 1 617 638 6381
Email: belindab@bu.edu

Abstract

Background: Both mHealth and eHealth interventions for smoking cessation are rapidly being developed and tested. There are no data on use of mHealth and eHealth technologies by smokers in general or by smokers who are not motivated to quit smoking.

Objective: The aims of our study were to (1) assess technology use (eg, texting, social media, Internet) among smokers in the United States and United Kingdom, (2) examine whether technology use differs between smokers who are motivated to quit and smokers who are not motivated to quit, (3) examine previous use of technology to assist with smoking cessation, and (4) examine future intentions to use technology to assist with smoking cessation.

Methods: Participants were 1000 adult smokers (54.90%, 549/1000 female; mean age 43.9, SD 15.5 years; US: n=500, UK: n=500) who were recruited via online representative sampling strategies. Data were collected online and included demographics, smoking history, and frequency and patterns of technology use.

Results: Among smokers in general, there was a high prevalence of mobile and smartphone ownership, sending and receiving texts, downloading and using apps, using Facebook, and visiting health-related websites. Smokers who were unmotivated to quit were significantly less likely to own a smartphone or handheld device that connects to the Internet than smokers motivated to quit. There was a significantly lower prevalence of sending text messages among US smokers unmotivated to quit (78.2%, 179/229) versus smokers motivated to quit (95.0%, 229/241), but no significant differences between the UK groups (motivated: 96.4%, 239/248; unmotivated: 94.9%, 223/235). Smokers unmotivated to quit in both countries were significantly less likely to use a handheld device to read email, play games, browse the Web, or visit health-related websites versus smokers motivated to quit. US smokers had a high prevalence of app downloads regardless of motivation to quit, but UK smokers who were motivated to quit had greater prevalence of app downloads than smokers unmotivated to quit. US smokers were significantly more likely to have a Facebook account (87.0%, 435/500) than UK smokers (76.4%, 382/500), but smokers unmotivated to quit in both countries used Facebook less frequently than smokers motivated to quit. Smokers who were unmotivated to quit were less likely to have used eHealth or mHealth platforms to help them quit smoking in the past and less likely to say that they would use them for smoking cessation in the future.

Conclusions: Although smokers unmotivated to quit make less use of technology than smokers motivated to quit, there is sufficient prevalence to make it worthwhile to develop eHealth and mHealth interventions to encourage cessation. Short and low-effort communications, such as text messaging, might be better for smokers who are less motivated to quit. Multiple channels may be required to reach unmotivated smokers.

(J Med Internet Res 2015;17(7):e164) doi:10.2196/jmir.4420

XSL•FO
RenderX

KEYWORDS

smoking cessation; eHealth; mHealth; health behavior; motivation; text messaging

Introduction

The current prevalence of cigarette smoking is 18.1% in the United States [1] and 19% in the United Kingdom [2], with substantially higher prevalence among specific underserved subpopulations [3,4]. Although rates of smoking have declined over the last 10 years, this decline has recently plateaued [1,2], possibly due to lower cessation among specific subpopulations, such as those with low income and education and those with medical and psychiatric comorbidities [4]. In the United States and United Kingdom, national studies have shown that 29% to 31% of smokers are not interested in quitting smoking in either the short or long term [5,6].

The use of technology-based interventions, such as those delivered through Internet (eHealth) and mobile phones (mHealth), may enhance the reach of smoking cessation interventions given the lack of disparities by race, education, and income in use of these technologies [7-10]. For example, rates of mobile phone penetration are expected to reach close to 100% [11]; 81% of mobile phone owners use their phone to send or receive text messages [12] and the average number of text messages per day is high across all racial and ethnic groups (black: mean 70.1, median 20; Hispanic: mean 48.9, median 20; white: mean 31.2, median 10 [8]). Moreover, those who are Latino, black, or aged between 18 and 49 years are more likely to gather health information through their mobile phones [7].

Both mHealth and eHealth interventions have been shown to be effective for smoking cessation among those who are ready to quit [13-16]. However, mHealth and eHealth have not yet been used to motivate quit attempts in smokers who are not motivated to quit. Given the extent to which technology has become integrated into people's everyday lives, targeting smokers who are not motivated to quit through these platforms may help to jump-start stalled smoking cessation rates. To our knowledge, there are no studies that assess mHealth and eHealth use (and frequency of use) by smokers in general or by smokers who are not motivated to quit smoking. Before the development of mHealth and eHealth interventions for these smokers, it is necessary to ascertain their level of engagement with technology. This information could help intervention planners and funders to find out where the smokers "hang out" and direct resources accordingly. Thus, the aims of our study are to (1) assess technology use (eg, texting, social media, Internet) among smokers in the United States and United Kingdom, (2) examine whether use of technology differs between smokers who are motivated to quit and smokers who are not motivated to quit, (3) examine previous use of technology to assist with smoking cessation, and (4) assess future intentions to use technology to assist with smoking cessation.

We assessed smokers in the United States and United Kingdom because they are the 2 English-speaking markets with the most active users of iOS (iPhone/iPad) and Android devices [17]. Given that smokers who are not ready to quit comprise a large minority of the smoking population, understanding their use of, and level of engagement with, technology could help expand the reach of current smoking cessation interventions and develop interventions specifically targeted to, and tailored for, this population.

Methods

Participants

Participants were 1000 current smokers: 500 in the United States and 500 in the United Kingdom. In each country, we recruited 250 smokers who did not want to quit smoking (defined as "does not plan to quit smoking cigarettes in the next 30 days") and 250 smokers who were ready to quit smoking within 30 days and were either (1) currently investigating options for help with quitting smoking or (2) had set a quit date within 30 days. Participants were eligible if they were current, regular smokers (ie, smoke at least 3 tobacco cigarettes per day for the past year and smoked more than 100 cigarettes in their lifetime) and aged 18 years or older.

A total of 1767 people completed the initial screening questions; 572 were screened out because they did not meet eligibility criteria, 32 were removed due to random responding (see Data Analyses), and 89 were removed to ensure that the sample was representative of age and gender. Of those who were eligible to participate (n=1074), one participant was not able to be categorized as "motivated to quit" or "unmotivated to quit" and 6.80% (73/1074) did not complete the survey. Thus, the final sample was 1000 smokers: 500 in the United States and 500 in the United Kingdom.

Procedure

Participants were recruited through online survey sampling conducted by Toluna, Inc. Toluna has processes in place to ensure that respondents do not misrepresent themselves to gain access to a study for which they are not eligible and that no participant takes part in any study more than once. Participants were recruited from Toluna's panel, Toluna-affiliated partnerships, websites, and social media. All potential participants were extensively verified and underwent checks to ascertain their identity and location. Toluna also checked for duplication within the panel before permitting access to the survey. Participants received 4000 "panel points" for survey completion. These points could be redeemed for vouchers for shops and services, redeemed as cash, or used to enter prize drawings at the participant's discretion.

All data were collected during one week in August 2014 and 21.90% (219/1000) of the sample completed the survey on their mobile phone. Toluna removed all identifiable information before transferring the dataset to investigators (ie, removal of IP addresses). Toluna adheres to and exceeds various data security protocols regarding personal identifiable information for its panelists and its research respondents and they meet all international data security protocols (eg, ISO27001). Ethical approval was obtained from The Miriam Hospital in the United States and the University of Manchester in the United Kingdom,

XSL•FO
RenderX

and informed consent was obtained from participants before participation.

Measures

Demographics and smoking history were assessed with age, gender, marital status, ethnicity, employment, years of education, and number of cigarettes smoked per day. We assessed nicotine dependence with one item from the Fagerström Test for Nicotine Dependence (FTND) [18]: "Do you smoke within 30 minutes of waking? (yes/no)." This single item is highly correlated with the full scale [19]. General use of technology was assessed with the Technology Use Questionnaire, a series of questions developed for this study regarding use and frequency of use of different types of Internet and mobile technologies. We also used the social media and smartphone usage subscales from the Media and Technology Usage and Attitudes Scale [20]. We defined "regular users" as those who used a feature several times per month or once per day or more. The Technology-Assisted Smoking Cessation Questionnaire, also developed for this study, was used to assess the use of technology for smoking cessation. It consisted of 5 items asking about previous use and 5 items asking about future intentions to use each of the following technologies for smoking cessation: the Internet, text messages, mobile phone apps, Twitter, or Facebook.

Data Analysis

Data were cleaned before analyses. We eliminated (1) straight-liners (n=8), defined as respondents who selected the same answer option for all items within a scale so that they completed the survey as quick as possible with minimum effort and (2) speeders (n=24), who did not carefully read the questions and provided random responses as evidenced by completing the survey more quickly than the typical respondent (ie, completing the survey in less than half the median time). Toluna also checked for respondents who filled in random letters into open-ended question fields, those who inserted offensive words, and duplicate survey takers, but none were noted in our sample. The final sample size was 1000.

We analyzed the overall prevalence of technology use and frequency of use by smokers, as well as compared differences between smokers who were motivated to quit and smokers who were not motivated to quit. Independent group *t* tests were computed for continuous variables and chi-square tests were computed for categorical data. For chi-square tests, standardized residuals were calculated and values equal to or greater than 1.96 or equal to or less than -1.96 (the critical value that corresponds to $\alpha<.05$) were considered large (ie, [21]). We used hierarchical logistic regression to ascertain whether motivation group (motivated to quit vs unmotivated to quit) was related to device ownership and technology use after controlling for demographics (age, education, ethnicity, and income).

Results

Sample Demographics

Only 6.80% (73/1074) of the sample did not complete the survey. Noncompleters (n=73) were significantly more likely to be female ($\chi^2_1=6.2$, $P=.01$) and were more likely to complete the survey on a mobile device ($\chi^2_1=7.3$, $P<.007$). There were no differences between completers and noncompleters in age, country, group (motivated vs unmotivated), number of cigarettes smoked per day, when they planned to quit smoking, and confidence in their ability to quit smoking.

The sample was comprised of 54.90% (549/1000) female smokers (Table 1) and the ethnic composition was 82.60% (826/1000) white, 6.50% (65/1000) black, 3.70% (37/1000) Hispanic/white, 0.40% (4/1000) Hispanic/black, 3.10% (31/1000) Asian, 0.40% (4/1000) American Indian/Alaskan Native, 0.30% (3/1000) Native Hawaiian or Pacific islander, 1.50% (15/1000) multiracial, and 1.50% (15/1000/) prefer not to say. Slightly more than half (55.90%, 559/1000) of the sample was employed (full- or part-time) and 61.30% (613/1000) were partnered (ie, married, engaged, living together, or in relationship but not living together). Participants smoked mean 16.5 (SD 13.4) cigarettes per day and 72.30% (723/1000) smoked within 30 minutes of waking, suggesting a high level of behavioral dependence on nicotine [19].

We assessed demographic differences between smokers who were motivated to quit and smokers who were not motivated to quit. Smokers who were not motivated to quit were significantly older ($t_{998}=14.31$, $P<.001$) and less likely to be employed ($\chi^2_1=19.3$, $P<.001$) than smokers who were motivated to quit. Ethnic minorities were more likely to be motivated to quit (71.7%, 114/159) than white smokers (45.5%, 376/826; $\chi^2_1=36.6$, $P<.001$). There were no other significant demographic differences between motivated and unmotivated smokers (Table 1). In terms of differences between the samples in the 2 countries, the US sample had a significantly higher proportion of females ($\chi^2_1=7.5$, $P=.006$), were less likely to report paid employment ($\chi^2_1=7.5$, $P=.006$), and were significantly more likely to complete the survey on a mobile phone (26.8%, 134/500) than UK participants (17.0%, 85/500; $\chi^2_1=14.0$, $P<.001$).

XSL•FO
RenderX

Table 1. Demographics of the total sample and by motivation to quit.

Variable	Total sample N=1000	Smokers not motivated to quit n=500	Smokers motivated to quit n=500	χ^2_1	t_{998}	*P*
Female, n (%)	549 (54.90)	286 (52.1)	263(47.9)	2.1		.14
Age (years), mean (SD)	43.9 (15.4)	50.3 (14.1)	37.5 (14.1)		14.31	<.001
Ethnicity (white), n (%)	826 (82.60)	450 (54.5%)	376 (45.5)	36.6		<.001
<University education,[a] n (%)	488 (49.10)	253 (51.8)	235 (48.2)	1.10		.30
Employed full- or part-time, n (%)	559 (55.90)	245 (43.8)	314 (56.2)	19.3		<.001
Partnered/in relationship, n (%)	613 (61.30)	312 (50.9)	301 (49.1)	0.5		.48
Cigarettes smoked/day, mean (SD)	16.6 (13.4)	17.3 (12.6)	15.8 (14.1)		1.76	.08

[a] Participants selecting “I don’t know” were counted as missing.

Smokers in the United States

Prevalence of Handheld Device Ownership and Differences by Motivation to Quit

Of the US sample, 92.8% (464/500) reported owning a mobile phone and 75.9% (352/464) of these were smartphones (Table 2). In addition, 79.6% (374/470) reported that they owned a handheld device that connects to the Internet and 49.8% (249/500) reported owning a tablet. Smokers who were motivated to quit were significantly more likely to own a mobile device, such as a mobile phone or tablet (95.6%, 239/250), than smokers who were not motivated to quit (90.0%, 225/250). Of those who owned devices, smokers who were motivated to quit (86.3%, 208/500) were more likely to be able to access the Internet on their mobile phone or tablet versus smokers not motivated to quit (72.5%, 166/500). There were no differences between groups after controlling for demographics in regression analyses.

Prevalence of Types of Technology and Frequency of Use

Of those who had devices capable of text messaging, only 13.2% (62/470) reported that they never send text messages and 10.6% (50/470) reported that they never receive text messages. Of those who sent text messages, 32.8% (134/408) sent 2 to 9 texts per day and 34.1% (139/408) sent 10 or more texts per day (“supertexters”). Of those who received text messages, 33.5% (141/421) received 2 to 9 texts per day and 32.8% (138/421) received 10 or more texts per day.

Of those who reported having handheld devices (94.0% 470/500), the most common features regularly used (ie, several times per month or more) were reading email (68.3%, 321/470), browsing the Web (70%, 329/470), taking photos (66.0%, 310/470), using apps (66.6%, 313/470), and playing games (58.7%, 276/470). Additionally, 70% (350/500) of this sample had visited health-related websites on either a handheld device or computer; of these, only 34.0% (119/350) visited them regularly (twice per week or more). Of those who owned a handheld device capable of accessing the Internet, 91.4% (342/374) reported that they had previously downloaded an app; of those, 26.4% (n95/360) said that they used it for 1 month or more (but less than 1 year) and 34.4% (124/360) said that they used it for 1 year or more. Only 10.0% (36/360) reported that they downloaded an app, but used it for less than 1 day. For Facebook, 87% (435/500) reported having a Facebook account; 20.5% (89/435) checked it once per day and 49.7% (n216/435) checked it more than once per day.

Prevalence of Types of Technology and Frequency of Use: Differences by Motivation to Quit

There were differences in technology use and frequency of use between smokers who were motivated to quit and smokers who were not motivated to quit. Only 5.0% (12/241) of smokers who were motivated to quit reported never sending text messages versus 21.8% (50/229) of smokers who were not motivated to quit. Differences between motivation groups were maintained after demographics were controlled for in regression analysis (*P*=.01, Wald=6.08, SE=0.40). Of those who sent texts, there was a significant relationship between type of smoker (motivated vs unmotivated) and frequency of texting. Examination of the standardized residuals suggested that motivated smokers were more likely to be supertexters (texting >10 times per day).

Smokers who were motivated to quit were significantly more likely than smokers unmotivated to quit to regularly use their handheld devices to accomplish a variety of tasks (eg, email, browse the Web, use apps, play games; Table 2). Of smokers who were not motivated to quit, 43.2% (108/250) reported that they never visited websites related to health issues versus 16.8% (42/250) of smokers who were motivated to quit (differences that were maintained after controlling for demographics; *P*=.006, Wald=7.51, SE=0.25). Of those who visited health-related websites, smokers who were motivated to quit visited these websites more frequently than smokers unmotivated to quit.

The majority of smokers who were motivated to quit (93.8%, 195/208) reported that they previously downloaded an app and this prevalence was not significantly different from that of unmotivated smokers (88.6%, 147/166; *P*=.07). There were no significant differences between motivation groups in the longest length of time of app use. Although there were no significant differences between motivated and unmotivated smokers in the prevalence of having a Facebook account (*P*=.14), smokers who were motivated to quit checked Facebook more frequently (>once per day: 57.8%, 129/223) than smokers who were unmotivated to quit (>once per day: 41.0%, 216/435).

XSL•FO
RenderX

Table 2. Prevalence of technology use among smokers in the United States and differences in technology use by motivation to quit.

Variable	Total US sample, n (%) n=500	Smokers not motivated to quit, n (%) n=250	Smokers motivated to quit, n (%) n=250	χ^2 (*df*)	*P*
Device ownership					
Own a mobile	464 (92.8)	225 (90.0)	239 (95.6)	5.9 (1)	.02
Own a tablet	249 (49.8)	100 (40.0)	149 (59.6)	19.2 (1)	<.001
Mobile is a smartphone	352 (75.9)	153 (68.0)	199 (83.3)	14.7 (1)	<.001
Internet-enabled handheld device	374 (79.6)	166 (72.5)	208 (86.3)	13.8 (1)	<.001
Texting					
Send text messages	408 (86.8)	179 (78.2)	229 (95.0)	29.1 (1)	<.001
Receive text messages	421 (89.6)	190 (83.0)	231 (95.9)	20.9 (1)	<.001
Frequency of sending texts [a]				33.4 (5)	<.001
≤1 texts per month	31 (7.6)	14 (7.8)	17 (7.4)		
2-4 texts per month[b]	29 (7.1)	22 (12.3)	7 (3.1)		
2-6 texts per week[c]	56 (13.7)	35 (19.6)	21 (9.2)		
1 text per day	19 (4.7)	10 (5.6)	9 (3.9)		
2-9 texts per day	134 (32.8)	58 (32.4)	76 (33.2)		
≥10 texts per day[b]	139 (34.1)	40 (22.3)	99 (43.2)		
Frequency of receiving texts [d]				32.4 (5)	<.001
≤1texts per month	34 (8.1)	17 (8.9)	17 (7.4)		
2-4 texts per month[c]	36 (8.6)	26 (13.7)	10 (4.3)		
2-6 texts per week[c]	55 (13.1)	36 (18.9)	19 (8.2)		
1 text per day	17 (4.0)	8 (4.2)	9 (3.9)		
2-9 texts per day	141 (33.5)	61 (32.1)	80 (34.6)		
≥10 texts per day[b]	138 (32.8)	42 (22.1)	96 (41.6)		
Features of a handheld device used regularly					
Read email	321 (68.3)	137 (59.8)	184 (76.3)	14.8 (1)	<.001
Get directions or use navigation (eg, GPS)	226 (48.1)	83 (36.2)	143 (59.3)	25.1 (1)	<.001
Browse the Web	329 (70.0)	136 (59.4)	193 (80.1)	24.0 (1)	<.001
Listen to music	276 (58.7)	103 (45.0)	173 (71.8)	34.8 (1)	<.001
Take photos	310 (66.0)	121 (52.8)	189 (78.4)	34.2 (1)	<.001
Check the news	270 (57.4)	106 (46.3)	164 (68.0)	22.7 (1)	<.001
Record video	168 (35.7)	57 (24.9)	111 (46.1)	22.9 (1)	<.001
Use apps (for any purpose)	313 (66.6)	130 (56.8)	183 (75.9)	19.4 (1)	<.001
Search for information	315 (67.0)	127 (55.5)	188 (78.0)	27.0 (1)	<.001
Play games by yourself	276 (58.7)	114 (49.8)	162 (67.2)	14.7 (1)	<.001
Play games with other people	168 (35.7)	46 (20.1)	122 (50.6)	47.7 (1)	<.001
Apps					
Previous app download[e]	342 (91.4)	147 (88.6)	195 (93.8)	3.2 (1)	.07
Longest used app for[e]				7.1 (4)	.13
<1 day	36 (10.0)	13 (8.6)	23 (11.1)		
≥1 day but <1 week	56 (15.6)	21 (13.8)	35 (16.8)		

XSL•FO
RenderX

Variable	Total US sample, n (%) n=500	Smokers not motivated to quit, n (%) n=250	Smokers motivated to quit, n (%) n=250	χ^2 (*df*)	*P*
≥1 week but <1 month	49 (13.6)	16 (10.5)	33 (15.9)		
≥1 month but <1 year	95 (26.4)	50 (32.9)	45 (21.6)		
≥1 year	124 (34.4)	52 (34.2)	72 (34.6)		
Facebook					
Facebook account	435 (87.0)	212 (84.8)	223 (89.2)	2.1 (1)	.14
Frequency of checking [f]				15.7 (5)	.008
Never	8 (1.8)	3 (1.4)	5 (2.2)		
≤Once per month	24 (5.5)	17 (8.0)	7 (3.1)		
2-4 times per month	26 (6.0)	15 (7.1)	11 (4.9)		
2-6 times per week	72 (16.6)	42 (19.8)	30 (13.5)		
Once per day	89 (20.5)	48 (22.6)	41 (18.4)		
>Once per day	216 (49.7)	87 (41.0)	129 (57.8)		
Visit health-related websites	350 (70.0)	142 (56.8)	208 (83.2)	41.5 (1)	<.001
Frequency of visits [g]				38.5 (4)	<.001
≤Once per month[b]	125 (35.7)	71 (50.0)	54 (26.0)		
2-4 times per month	106 (30.3)	45 (31.7)	61 (29.3)		
2-6 times per week	54 (15.4)	18 (12.7)	36 (17.3)		
Once per day	34 (9.7)	7 (4.9)	27 (13.0)		
>Once per day[b]	31 (8.9)	1 (0.7)	30 (14.4)		

[a] Of those who sent text messages (n=408).

[b] Standardized residual ≥2.58 or ≤–2.58.

[c] Standardized residual ≥1.96 or ≤–1.96

[d] Of those who receive text messages (n=421).

[e] Of those who can access the Internet on their handheld device (n=374).

[f] Of those who have a Facebook account (n=435).

[g] Of those who visit websites related to health issues (n=350).

Smokers in the United Kingdom

Prevalence of Handheld Device Ownership and Differences by Motivation to Quit

Of the UK sample, 95.8% (479/500) reported owning a mobile phone and 82.9% (397/479) of these were smartphones (Table 3). In all, 85.7% (414/483) reported that they owned a handheld device that connects to the Internet and 56.2% (281/500) reported owning a tablet. Smokers who were motivated to quit were significantly more likely to own a handheld device (mobile or tablet) than smokers unmotivated to quit (99.2%, 248/250 vs 94.0%, 235/250). Of those who owned handheld devices, smokers who were motivated to quit (90.7%, 225/248) were significantly more likely to have a device that connects to the Internet than smokers who were not motivated to quit (80.4%, 189/235). There were no differences between motivation groups after controlling for demographics (age, education, and income) in logistic regressions.

Prevalence of Types of Technology and Frequency of Use Among UK Smokers

The vast majority of UK smokers reported that they send (95.7%, 462/483) and receive (97.7%, 472/483) text messages. Of those who send text messages, 36.1% (167/462) send 2 to 9 texts per day and 21.4% (99/462) were supertexters. Of those who received text messages, 36.7% (173/472) received 2 to 9 texts per day and 21.2% (100/472) received 10 or more texts per day.

Of those who had handheld devices (96.6%, 483/500), the most common features regularly used (ie, several times per month or more) were browsing the Web (70.2%, 339/483), reading email (70.0%, 338/483), searching for information (68.9%, 333/483), using apps (66.5%, 321/483), and taking photos (62.1%, 300/483). Of those who ever visited health-related websites (61.0%, 305/500), 24.9% (76/305) visited them regularly (twice per week or more). Of those with Internet-enabled handheld devices (85.7%, 414/483), 86.2% (357/414) reported that they had previously downloaded an app; 26.6% (95/357) said that they used it for 1 month or more (but less than 1 year) and 34.2%

XSL•FO
RenderX

(122/357) said that they used it for 1 year or more. Only 10.1% (36/357) said that they downloaded an app but used it less than 1 day; 29.1% (104/357) used it more than 1 day but less than 1 month. Of the UK sample, 76.4% (382/500) reported having a Facebook account; 20.7% (79/382) checked it once per day and 53.7% (205/382) checked it more than once per day.

Prevalence of Types of Technology and Frequency of Use: Differences by Motivation to Quit

There were no significant differences between smokers who were motivated to quit and smokers who were not motivated to quit in whether or not they texted; however, the prevalence of texting was very high among both groups: only 3.6% (9/248) of smokers who were motivated to quit reported never sending text messages versus 5.1% (12/235) of smokers who were not motivated to quit (P=.43). However, of those who texted, there was a significant relationship between motivation to quit and frequency of texting. Examination of the standardized residuals indicated that smokers who were not motivated to quit were less likely to reside in the supertexter category (eg, >10 texts per day). Smokers who were motivated to quit were significantly more likely to regularly use their handheld devices for a variety of tasks (eg, email, browse the Web, use apps) than smokers who were not motivated to quit (Table 3). Of smokers who were not motivated to quit, 54.8% (137/250) reported that they never visited websites related to health issues versus 23.2% (58/250) of smokers who were motivated to quit; this difference was maintained when demographic covariates were controlled for in a logistic regression analysis (P<.001, Wald=26.92, SE=0.22). Of those who visited health-related websites, there was a significant relationship between motivation to quit and frequency of visits.

Among those who had Internet access on their handheld devices (85.7%, 414/483), smokers who were motivated to quit were significantly more likely to have previously downloaded an app (91.6%, 206/225) than smokers who were not motivated to quit (79.9%, 151/189). There were no significant group differences when demographic variables were controlled.

Smokers who were motivated to quit were significantly more likely to have a Facebook account (82.4%, 206/250) than smokers who were not motivated to quit (70.4%, 176/250). There were no significant differences between groups when demographic variables were controlled for in a logistic regression analysis. Of those who had a Facebook account, there was a significant relationship between motivation to quit and the frequency of checking Facebook. Although none of the standardized residuals were equal to or greater than 1.96 or equal to or less than –1.96, the largest difference in percentages showed that smokers who were motivated to quit were more likely to report checking their Facebook pages more than once per day than smokers who were not motivated to quit (58.3%, 120/206 vs 48.3%, 85/176).

XSL•FO
RenderX

Table 3. Prevalence of technology use among smokers in the United Kingdom and differences in technology use by motivation to quit.

Variable	Total UK sample, n (%) n=500	Smokers not motivated to quit, n (%) n=250	Smokers motivated to quit, n (%) n=250	χ^2 (*df*)	*P*
Device ownership					
Own a mobile	479 (95.8)	233 (93.2)	246 (98.4)	8.40 (1)	.004
Own a tablet	281 (56.2)	109 (43.6)	172 (68.8)	32.3 (1)	<.001
Mobile is a smartphone	397 (82.9)	173 (74.2)	224 (91.1)	23.8 (1)	<.001
Handheld device connects to the Internet	414 (85.7)	189 (80.4)	225 (90.7)	10.5 (1)	.001
Texting					
Send text messages	462 (95.7)	223 (94.9)	239 (96.4)	0.6 (1)	.43
Receive text messages	472 (97.7)	232 (98.7)	240 (96.8)	2.1 (1)	.15
Frequency of sending texts [a]				27.7 (5)	.001
≤1 texts per month	41 (8.9)	27 (12.1)	14 (5.9)		
2-4 texts per month	38 (8.2)	25 (11.2)	13 (5.4)		
2-6 texts per week	87 (18.8)	48 (21.5)	39 (16.3)		
1 text per day	30 (6.5)	17 (7.6)	13 (5.4)		
2-9 texts per day	167 (36.1)	72 (32.3)	95 (39.7)		
>10 texts per day[b]	99 (21.4)	34 (15.2)	65 (27.2)		
Frequency of receiving texts [c]				26.3 (5)	<.001
≤1 texts per month	32 (6.8)	22 (9.5)	10 (4.2)		
2-4 texts per month[b]	48 (10.2)	34 (14.7)	14 (5.8)		
2-6 texts per week	84 (17.8)	44 (19.0)	40 (16.7)		
1 text per day	35 (7.4)	16 (6.9)	19 (7.9)		
2-9 texts per day	173 (36.7)	84 (36.2)	89 (37.1)		
>10 texts per day[b]	100 (21.2)	32 (13.8)	68 (28.3)		
Features of a handheld device used regularly					
Read email	338 (70.0)	139 (59.1)	199 (80.2)	25.6 (1)	<.001
Get directions or use navigation (eg, GPS)	184 (38.1)	60 (25.5)	124 (50.0)	30.6 (1)	<.001
Browse the Web	339 (70.2)	141 (60.0)	198 (79.8)	22.7 (1)	<.001
Listen to music	257 (53.2)	88 (37.4)	169 (68.1)	45.7 (1)	<.001
Take photos	300 (62.1)	118 (50.2)	182 (73.4)	27.5 (1)	<.001
Check the news	286 (59.2)	111 (47.2)	175 (70.6)	27.2 (1)	<.001
Record video	135 (28.0)	35 (14.9)	100 (40.3)	38.8 (1)	<.001
Use apps (for any purpose)	321 (66.5)	129 (54.9)	192 (77.4)	27.5 (1)	<.001
Search for information	333 (68.9)	131 (55.7)	202 (81.5)	37.2 (1)	<.001
Play games by yourself	254 (52.6)	92 (39.1)	162 (65.3)	33.1 (1)	<.001
Play games with other people	121 (25.1)	34 (14.5)	87 (35.1)	27.3 (1)	<.001
Apps					
Previous app download[d]	357 (86.2)	151 (79.9)	206 (91.6)	11.8 (1)	.001
Longest used app for [d]				6.8 (4)	.14
<1 day	36 (10.1)	13 (8.6)	23 (11.2)		

XSL•FO
RenderX

Variable	Total UK sample, n (%) n=500	Smokers not motivated to quit, n (%) n=250	Smokers motivated to quit, n (%) n=250	χ^2 (*df*)	*P*
≥1day but <1 week	56 (15.7)	21 (13.9)	35 (17.0)		
≥1 week but <1 month	48 (13.4)	16 (10.6)	32 (15.5)		
≥1 month but <1 year	95 (26.6)	50 (33.1)	45 (21.8)		
≥1 year	122 (34.2)	51 (33.8)	71 (34.5)		
Facebook					
Facebook account	382 (76.4)	176 (70.4)	206 (82.4)	10.00 (1)	.002
Frequency of checking [e]				13.0 (5)	.02
Never	7 (1.8)	5 (2.8)	2 (1.0)		
≤Once per month	22 (5.8)	16 (9.1)	6 (2.9)		
2-4 times per month	26 (6.8)	12 (6.8)	14 (6.8)		
2-6 times per week	43 (11.3)	25 (14.2)	18 (8.7)		
Once per day	79 (20.7)	33 (18.8)	46 (22.3)		
>Once per day	205 (53.7)	85 (48.3)	120 (58.3)		
Visit health-related websites	305 (61.0)	113 (45.2)	192 (76.8)	52.5 (1)	<.001
Frequency of visits [f]				15.7 (4)	.004
≤Once per month	153 (50.2)	69 (61.1)	84 (43.8)		
2-4 times per month	76 (24.9)	28 (24.8)	48 (25.0)		
2-6 times per week	48 (15.7)	10 (8.8)	38 (19.8)		
Once per day	18 (5.9)	6 (5.3)	12 (6.3)		
>Once per day	10 (3.3)	0 (0)	10 (5.2)		

[a] Of those who sent text messages (n=462).

[b] Standardized residual ≥1.96 or ≤–1.96.

[c] Of those who receive text messages (n=472).

[d] Of those who could access the Internet on their handheld devices (n=414).

[e] Of those with a Facebook account (n=382).

[f] Of those who visit websites related to health issues (n=305).

Previous use of Technology-Assisted Smoking Cessation Among US and UK Smokers and Differences by Motivation to Quit

Overview

Approximately one-quarter of smokers in the United States and United Kingdom reported that they previously used the Internet to quit smoking but the use of other technologies was low, ranging from 7.6% (38/500) for Twitter use in the United Kingdom to 15.2% (76/500) for smoking cessation app use in the United States (Table 4). Smokers in the United States were significantly more likely to have previously used Twitter (χ^2_1=5.9, *P*=.02), text messaging (χ^2_1=4.9, *P*=.03), and Facebook (χ^2_1=4.4, *P*=.04) to help them quit smoking than smokers in the United Kingdom.

There were significant differences in Internet-assisted cessation between smokers who were motivated to quit and smokers unmotivated to quit. Smokers who were motivated to quit were significantly more likely to have previously used each of the 5 assessed technologies to help them quit than smokers unmotivated to quit and these differences between groups were maintained when demographic covariates were controlled (Internet: *P*<.001, Wald=47.75, SE=0.01; text: *P*=.002, Wald=9.54, SE=0.01; quit smoking app: *P*<.001, Wald=33.72, SE=0.01; Twitter: *P*<.001, Wald=12.26, SE=0.01; and Facebook: *P*<.001, Wald=20.46, SE=0.01).

Future Intentions to Use Technology-Assisted Smoking Cessation Among US and UK Smokers and Differences by Motivation to Quit

Across both countries, the platforms with the greatest percentage of people endorsing that they would use it to quit smoking in the future were the Internet (46.7%, 467/1000) and apps (42.7%, 427/1000) (Table 5). Smokers in the United States were more likely to report that they would use text messages (χ^2_1=4.1, *P*=.04) and Twitter (χ^2_1=6.1, *P*=.01) to quit smoking in the future than smokers in the United Kingdom. Smokers who were

XSL•FO
RenderX

motivated to quit were significantly more likely to say that they would use each of the 5 assessed technologies to help them quit smoking in the future than smokers unmotivated to quit, and these differences between groups were maintained when controlling for demographic covariates: future use of the Internet: $P<.001$, Wald=52.23, SE=0.01; text messages: $P=.001$, Wald=15.35, SE=0.01; quit smoking app: $P<.001$, Wald=54.40, SE=0.01; Twitter: $P<.001$, Wald=26.19, $SE=0.01$; and Facebook: $P<.001$, Wald=27.21, SE=0.01.

Table 4. Previous use of technology-assisted smoking cessation among smokers in the United States and United Kingdom and differences by motivation to quit.

Previous use of technology to quit smoking	Total sample, n (%)	Unmotivated smokers, n (%)	Motivated smokers, n (%)	χ^2_1	*P*
US Smokers					
Used the Internet (a website)	131 (26.2)	27 (10.8)	104 (41.6)	61.3	<.001
Joined a quit smoking program that involved text messaging	66 (13.2)	12 (4.8)	54 (21.6)	30.8	<.001
Used a quit smoking app on your phone	76 (15.2)	16 (6.4)	60 (24.0)	30.0	<.001
Used Twitter to connect with other smokers who are trying to quit	61 (12.2)	13 (5.2)	48 (19.2)	22.9	<.001
Used Facebook to connect with other smokers who are trying to quit	74 (14.8)	16 (6.4)	58 (23.2)	28.0	<.001
UK Smokers					
Used the Internet (a website)	128 (25.6)	23 (9.2)	105 (42.0)	70.6	<.001
Joined a quit smoking program that involved text messaging	44 (8.8)	6 (2.4)	38 (15.2)	25.5	<.001
Used a quit smoking app on your phone	64 (12.8)	7 (2.8)	57 (22.8)	44.8	<.001
Used Twitter to connect with other smokers who are trying to quit	38 (7.6)	3 (1.2)	35 (14.0)	29.2	<.001
Used Facebook to connect with other smokers who are trying to quit	52 (10.4)	8 (3.2)	44 (17.6)	27.8	<.001

Table 5. Future intentions to use technology-assisted smoking cessation among smokers in the United States and United Kingdom and differences by motivation to quit.

Future intentions to use technology to quit	Total sample, n (%)	Unmotivated smokers, n (%)	Motivated smokers, n (%)	χ^2_1	*P*
US Smokers					
Use the Internet (a website)	227 (45.4)	87 (34.8)	140 (56.0)	22.7	<.001
Join a quit smoking program that involves text messaging	157 (31.4)	56 (22.4)	101 (40.4)	18.8	<.001
Use a quit smoking app on your phone	217 (43.4)	90 (36.0)	127 (50.8)	11.2	<.001
Use Twitter to connect with other smokers who are trying to quit	113 (22.6)	34 (13.6)	79 (31.6)	23.2	<.001
Use Facebook to connect with other smokers who are trying to quit	152 (30.4)	53 (21.2)	99 (39.6)	20.0	<.001
UK Smokers					
Use the Internet (a website)	240 (48.0)	89 (35.6)	151 (60.4)	30.8	<.001
Join a quit smoking program that involves text messaging	128 (25.6)	43 (17.2)	85 (34.0)	18.5	<.001
Use a quit smoking app on your phone	210 (42.0)	76 (30.4)	134 (53.6)	27.6	<.001
Use Twitter to connect with other smokers who are trying to quit	82 (16.4)	22 (8.8)	60 (24.0)	21.1	<.001
Use Facebook to connect with other smokers who are trying to quit	129 (25.8)	40 (16.0)	89 (35.6)	25.1	<.001

XSL•FO
RenderX

Discussion

Principal Findings

Both mHealth and eHealth interventions for smoking cessation are rapidly being developed and tested, but to our knowledge, there are no data on use of these technologies by smokers in general or whether use differs by motivation to quit. Knowing the types of technologies that smokers engage with can help intervention planners design interventions that target smokers more effectively and efficiently. The aims of our study were to (1) assess technology use among smokers in the United States and United Kingdom, (2) examine whether technology use differs between smokers who are motivated to quit and smokers who are not motivated to quit, (3) examine previous use of technology-based assisted smoking cessation, and (4) examine future intentions to use technology-based assisted smoking cessation. The advantages of mobile platforms include the ability to implement interventions in real time and access them any time and from any place, ability to tailor to user needs (eg, content, timing, and intensity), few barriers to participation, decreased time gap between treatment and behavior, low participant burden (particularly important for smokers who are less motivated to quit), ability to provide instant support, ability to provide feedback on goal setting and achievement, capability for integration with social networking, and scalability to large populations.

Among smokers in general, we found a high prevalence of mobile and smartphone ownership, sending and receiving texts, downloading and using apps, using Facebook, and visiting websites related to health. The use of these platforms, however, has outpaced the ability to gather scientific evidence regarding their effectiveness for smoking cessation. Although more than 400 smoking cessation mobile apps were available in 2013 [22], no fully powered randomized trials regarding their efficacy have been published yet (for pilot trials see [23,24]). Meta-analyses have shown that text messaging for smoking cessation is promising [13,25], but studies published to date have suffered from one or more methodological shortcomings, including lack of power, nonrandomization, short-term follow-up, lack of adequate control groups, lack of biochemical verification of cessation, and experimental designs that make it difficult to isolate the effects of text messaging on smoking cessation (eg, multicomponent interventions). With regard to Facebook, although 3 trials are currently underway, two do not assess smoking cessation as an outcome [26,27] and one is recruiting only smokers aged between 18 and 25 years [28]. Studies are also needed on dosage (eg, number of texts or app notifications needed for effectiveness), features that have the greatest potency for changing smoking behavior, theoretical mechanisms of action (“why” it works), and effectiveness for special populations (eg, those who are not motivated to quit).

The second aim of our study was to evaluate whether use of these platforms differs between smokers who are motivated to quit and those who are not motivated to quit. Although the prevalence of texting was high in both countries, US smokers who were not motivated to quit were less likely to text than US smokers who were motivated to quit. In both countries, smokers who were motivated to quit were supertexters and unmotivated smokers tended to text less frequently. More than a quarter of our total sample and approximately one-fifth of unmotivated smokers said they would be willing to use a text message program to quit smoking in the future. The ubiquity and frequency of text messaging among smokers in general and among unmotivated smokers specifically lends support to the idea that text messaging could also be used to motivate smokers to quit, perhaps serving as a way to keep cessation “on the radar,” titrating the number of text messages upward when and if the smoker becomes motivated to quit. Creative ways to keep unmotivated smokers engaged with this process should be explored. To date, text message interventions for smoking cessation have targeted only smokers who are ready and willing to set a quit date within 30 days.

Compared with unmotivated smokers, smokers who were motivated to quit tended to use their handheld devices more often to read email, get directions, browse the Web, listen to music, take photos/video, check the news, search for information, and play games. Intervention planners could capitalize on this information by examining the most prevalent features used by smokers and how they might reach smokers through these features. For example, more than 65% smokers who were motivated to quit reported that they played games on their handheld device. Thus, gaming principles could be incorporated into mobile cessation, possibly curbing smoking urges, providing distraction during times of temptation, and promoting self-efficacy for quitting. One preliminary study has shown that a prototype of an interactive game was engaging to smokers [29]. Our study extends this research by our finding that a large minority of unmotivated smokers play video games (39.2% in United Kingdom and 49.8% in United States), so involving gaming in promoting motivation to quit could be explored.

App downloads and length of app use differed by motivation group and by country. In the United Kingdom, motivated smokers were more likely to have downloaded an app than unmotivated smokers, but there were no differences in the length of time that apps were used. In the United States, there were no differences between motivation groups in the prevalence of downloading apps (both >88%). Research is needed regarding what features make a smoking cessation app “sticky” (ie, has staying power with the user) and designing apps that adhere to the human-centered design principles, including health literacy [30,31].

Although the vast majority of smokers had a Facebook account (UK: 76.4%; US: 87.0%), there were significant differences by motivation to quit for UK smokers, such that motivated smokers were more likely to have an account than unmotivated smokers. Regardless of motivation to quit, more than 72% checked Facebook once per day or more. There may be several advantages to delivering health behavior interventions through Facebook [32], such as the ability to reach participants while they interact in near real time, delivery of tailored content, ability to interface with apps, potential to influence perceptions of social norms [33], promote expansion of social networks beyond one’s own (which may have a high proportion of smokers), and incorporation of other media (eg, photos, video).

XSL•FO
RenderX

Social network analytics can be automatically collected and used to measure delivery and use of behavior change techniques. Research is underway on using Facebook as a self-propagating delivery channel for smoking cessation by influencing behavior in local networks and facilitating diffusion ("viral spread") between networks [26].

One striking finding is that unmotivated smokers were less likely to visit health-related websites (on their computer or handheld device) than were motivated smokers. Thus, these smokers need to be reached proactively through other media channels. This parallels the recommendations that were put forth before the advent of technology, that public health impact for smoking cessation could be achieved through proactive reach through existing infrastructures where smokers are located, such as primary care [34] and home health care [35], but also nontraditional settings such as beauty salons [36] and churches [37]. The current zeitgeist calls for finding and meeting smokers where they are located "electronically." Previous proactive reach involved building relationships between smokers and providers to "hook" the unmotivated smokers. The next challenge for eHealth and mHealth is to find the electronic hook, one that will engage smokers regardless of their motivation to quit.

We assessed previous use of technology-assisted smoking cessation and found that more than 25% of smokers in both countries used the Internet to quit smoking. Other technology platforms had very low prevalence (Table 4), ranging from 12% to 15% in the United States and 7.6% to 12.8% in the United Kingdom. Thus, there is a discrepancy between the prevalence of these platforms (eg, >85% of sample use text messaging) and the prevalence of using the platform for smoking cessation (eg, 13.2% of US smokers and 8.8% of UK smokers joined a text message program for smoking cessation). This discrepancy could be reflective of a dearth of text message programs, smokers' perception that text messaging would not be helpful, lack of marketing, or preference for human contact and support for smoking cessation. We also assessed future intentions to use technology-assisted smoking cessation and only two platforms (Internet and apps) exceeded 50% of the sample's endorsement and that occurred only among motivated smokers. Among unmotivated smokers, Internet and apps had the highest endorsement at just over one-third in both countries.

Limitations

One potential limitation of the current study is that participants were recruited online. There may be concern that this approach biases the sample to smokers who have Internet access. However, approximately 87% of the adult population of the United States and United Kingdom are Internet users [38,39], so our sample likely reflects the majority of smokers. In addition, we used a sampling strategy that ensured representativeness of smokers in each country. There may also be concern about the veracity of the data given that it was collected online. We checked for and eliminated duplicate IP addresses, random responders, and speeders and straight-liners to provide greater confidence in the data.

Conclusions

Smoking cessation is an important public health goal, but the rate of cessation appears to have plateaued, meaning that new approaches are required. One possible approach is to devote greater efforts to understanding ways to target smokers who are not currently thinking of quitting smoking. This paper shows that although smokers who are not currently thinking of quitting make less use of technology than do smokers who are motivated to quit, sufficient numbers do use technology to make it worthwhile to develop these technologies designed to encourage unmotivated smokers to quit. Examining how health behavior change programs can capitalize on high rates of technology use is a public health priority, particularly because of the lack of disparities in the use of these technologies, relative low cost [40,41], and high potential for both customization and scalability.

Acknowledgments

This work was funded by Cancer Research UK (Grant # C1005 A17906) and by internal funds from The Miriam Hospital while the first author was employed there. The sponsors did not have a role in the preparation or publication of this manuscript.

Conflicts of Interest

None declared.

References

1. Blackwell DL, Lucas JW, Clarke TC. Summary health statistics for US adults: National Health Interview Survey, 2012. Vital Health Stat 2014;10(260):1-171 [FREE Full text]
2. Office for National Statistics. 2013. Opinions and Lifestyle Survey, Smoking Habits Amongst Adults, 2012 URL: http://www.ons.gov.uk/ons/rel/ghs/opinions-and-lifestyle-survey/smoking-habits-amongst-adults--2012/rpt-opinions-and-lifestyle-survey---smoking-habits-amongst-adults--2012.html [accessed 2015-02-02] [WebCite Cache ID 6W2ofLu67]
3. Borrelli B. Smoking cessation: next steps for special populations research and innovative treatments. J Consult Clin Psychol 2010 Feb;78(1):1-12. [doi: 10.1037/a0018327] [Medline: 20099945]
4. Borrelli B, Busch A, Dunsiger S. Cigarette smoking among adults with mobility impairments: a US population-based survey. Am J Public Health 2014 Oct;104(10):1943-1949. [doi: 10.2105/AJPH.2013.301772] [Medline: 25208005]

XSL•FO
RenderX

5. Office for National Statistics. 2009. Smoking related behaviour and attitudes, 2008/09 URL: http://www.ons.gov.uk/ons/rel/lifestyles/smoking-related-behaviour-and-attitudes/2008-09/index.html [accessed 2015-02-02] [WebCite Cache ID 6W2oQTHsk]
6. Centers for Disease Control and Prevention. Quitting smoking among adults--United States, 2001-2010. MMWR Morb Mortal Wkly Rep 2011 Nov 11;60(44):1513-1519 [FREE Full text] [Medline: 22071589]
7. Fox S, Rainie L. The Web at 25 in the US. Washington, DC: Pew Internet & American Life Project; 2014 Feb 27. URL: http://www.pewinternet.org/2014/02/27/the-web-at-25-in-the-u-s/ [accessed 2015-02-02] [WebCite Cache ID 6W2oxeUUC]
8. Smith A. Americans and Text Messaging. Washington, DC: Pew Internet & American Life Project; 2011 Oct 19. URL: http://www.pewinternet.org/2011/09/19/americans-and-text-messaging/ [accessed 2015-02-02] [WebCite Cache ID 6W2tNijd2]
9. Pew Research Center. 2014. Data trend: cell phone and smartphone ownership demographics URL: http://www.pewinternet.org/data-trend/mobile/cell-phone-and-smartphone-ownership-demographics/ [accessed 2015-02-02] [WebCite Cache ID 6W2tPxmOB]
10. Kaplan RM, Stone AA. Bringing the laboratory and clinic to the community: mobile technologies for health promotion and disease prevention. Annu Rev Psychol 2013;64:471-498. [doi: 10.1146/annurev-psych-113011-143736] [Medline: 22994919]
11. International Telecommunication Union. 2014. The world in 2014: ICT facts & figures URL: http://www.itu.int/en/Pages/default.aspx [accessed 2015-02-02] [WebCite Cache ID 6Zia5VeXJ]
12. Duggan M. Cell Phone Activities 2013. Washington, DC: Pew Internet & American Life Project; 2013 Oct 19. URL: http://www.pewinternet.org/2013/09/19/cell-phone-activities-2013/ [accessed 2015-02-02] [WebCite Cache ID 6W2pcr5cr]
13. Whittaker R, McRobbie H, Bullen C, Borland R, Rodgers A, Gu Y. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2012;11:CD006611. [doi: 10.1002/14651858.CD006611.pub3] [Medline: 23152238]
14. Free C, Whittaker R, Knight R, Abramsky T, Rodgers A, Roberts IG. Txt2stop: a pilot randomised controlled trial of mobile phone-based smoking cessation support. Tob Control 2009 Apr;18(2):88-91. [doi: 10.1136/tc.2008.026146] [Medline: 19318534]
15. 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]
16. 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]
17. Farago P. Flurry Insights. 2013. China knocks off US to become world's top smart device market URL: http://www.flurry.com/bid/94352/China-Knocks-Off-U-S-to-Become-World-s-Top-Smart-Device-Market [accessed 2015-02-02] [WebCite Cache ID 6W2oIfdgc]
18. Heatherton TF, Kozlowski LT, Frecker RC, Fagerström KO. The Fagerström Test for Nicotine Dependence: a revision of the Fagerström Tolerance Questionnaire. Br J Addict 1991 Sep;86(9):1119-1127. [Medline: 1932883]
19. Baker TB, Piper ME, McCarthy DE, Bolt DM, Smith SS, Kim S, et al. Time to first cigarette in the morning as an index of ability to quit smoking: implications for nicotine dependence. Nicotine Tob Res 2007 Nov;9 Suppl 4:S555-S570 [FREE Full text] [doi: 10.1080/14622200701673480] [Medline: 18067032]
20. Rosen LD, Whaling K, Carrier LM, Cheever NA, Rokkum J. The Media and Technology Usage and Attitudes Scale: an empirical investigation. Comput Human Behav 2013 Nov 1;29(6):2501-2511 [FREE Full text] [Medline: 25722534]
21. Segars AH, Grover V. Re-examining perceived ease of use and usefulness: a confirmatory factor analysis. MIS Quarterly 1993 Dec;17(4):517. [doi: 10.2307/249590]
22. Abroms LC, Lee Westmaas J, Bontemps-Jones J, Ramani R, Mellerson J. A content analysis of popular smartphone apps for smoking cessation. Am J Prev Med 2013 Dec;45(6):732-736 [FREE Full text] [doi: 10.1016/j.amepre.2013.07.008] [Medline: 24237915]
23. Bricker JB, Mull KE, Kientz JA, Vilardaga R, Mercer LD, Akioka KJ, et al. Randomized, controlled pilot trial of a smartphone app for smoking cessation using acceptance and commitment therapy. Drug Alcohol Depend 2014 Oct 1;143:87-94. [doi: 10.1016/j.drugalcdep.2014.07.006] [Medline: 25085225]
24. Buller DB, Borland R, Bettinghaus EP, Shane JH, Zimmerman DE. Randomized trial of a smartphone mobile application compared to text messaging to support smoking cessation. Telemed J E Health 2014 Mar;20(3):206-214 [FREE Full text] [doi: 10.1089/tmj.2013.0169] [Medline: 24350804]
25. Whittaker R, Borland R, Bullen C, Lin RB, 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]
26. Cobb NK, Jacobs MA, Saul J, Wileyto EP, Graham AL. Diffusion of an evidence-based smoking cessation intervention through Facebook: a randomised controlled trial study protocol. BMJ Open 2014;4(1):e004089 [FREE Full text] [doi: 10.1136/bmjopen-2013-004089] [Medline: 24448847]
27. Sadasivam RS, Volz EM, Kinney RL, Rao SR, Houston TK. Share2Quit: Web-based peer-driven referrals for smoking cessation. JMIR Res Protoc 2013;2(2):e37 [FREE Full text] [doi: 10.2196/resprot.2786] [Medline: 24067329]

XSL•FO
RenderX

28. Ramo DE, Rodriguez TM, Chavez K, Sommer MJ, Prochaska JJ. Facebook recruitment of young adult smokers for a cessation trial: methods, metrics, and lessons learned. Internet Interv 2014 Apr;1(2):58-64. [doi: 10.1016/j.invent.2014.05.001] [Medline: 25045624]
29. Krebs P, Burkhalter JE, Snow B, Fiske J, Ostroff JS. Development and alpha testing of QuitIT: an interactive video game to enhance skills for coping with smoking urges. JMIR Res Protoc 2013;2(2):e35 [FREE Full text] [doi: 10.2196/resprot.2416] [Medline: 24025236]
30. Broderick J, Devine T, Langhans E, Lemerise AJ, Lier S, Harris L. Discussion Paper: Designing Health Literate Mobile Apps. Washington, DC: Institute of Medicine; 2014 Jan 28. URL: http://www.health.gov/communication/literacy/BPH-HealthLiterateApps.pdf [accessed 2015-03-03] [WebCite Cache ID 6Wl5tOM52]
31. De Vito Dabbs A, Myers BA, Mc Curry KR, Dunbar-Jacob J, Hawkins RP, Begey A, et al. User-centered design and interactive health technologies for patients. Comput Inform Nurs 2009;27(3):175-183 [FREE Full text] [doi: 10.1097/NCN.0b013e31819f7c7c] [Medline: 19411947]
32. Cobb NK, Graham AL, Byron MJ, Niaura RS, Abrams DB, Workshop Participants. Online social networks and smoking cessation: a scientific research agenda. J Med Internet Res 2011;13(4):e119 [FREE Full text] [doi: 10.2196/jmir.1911] [Medline: 22182518]
33. Morris ME, Consolvo S, Munson S, Patrick K, Tsai J, Kramer AD. Facebook for health: opportunities and challenges for driving behavior change. In: Proceedings of CHI EA '11 Extended Abstracts on Human Factors in Computing Systems. New York: ACM; 2011 Presented at: CHI EA '11; May 7-12, 2011; Vancouver, BC.
34. Goldstein MG, Niaura R, Willey C, Kazura A, Rakowski W, DePue J, et al. An academic detailing intervention to disseminate physician-delivered smoking cessation counseling: smoking cessation outcomes of the Physicians Counseling Smokers Project. Prev Med 2003 Feb;36(2):185-196. [Medline: 12590994]
35. Borrelli B, Novak S, Hecht J, Emmons K, Papandonatos G, Abrams D. Home health care nurses as a new channel for smoking cessation treatment: outcomes from project CARES (Community-nurse Assisted Research and Education on Smoking). Prev Med 2005;41(5-6):815-821. [doi: 10.1016/j.ypmed.2005.08.004] [Medline: 16182355]
36. Linnan LA, D'Angelo H, Harrington CB. A literature synthesis of health promotion research in salons and barbershops. Am J Prev Med 2014 Jul;47(1):77-85. [doi: 10.1016/j.amepre.2014.02.007] [Medline: 24768037]
37. Schorling JB, Roach J, Siegel M, Baturka N, Hunt DE, Guterbock TM, et al. A trial of church-based smoking cessation interventions for rural African Americans. Prev Med 1997;26(1):92-101. [doi: 10.1006/pmed.1996.9988] [Medline: 9010903]
38. Office for National Statistics. 2014. Internet access – households and individuals 2014 URL: http://www.ons.gov.uk/ons/rel/rdit2/internet-access---households-and-individuals/2014/stb-ia-2014.html [accessed 2015-02-02] [WebCite Cache ID 6W2otKnYM]
39. Pew Research Center. 2014. Data trend: Internet user demographics URL: http://www.pewinternet.org/data-trend/internet-use/latest-stats/ [accessed 2015-02-23] [WebCite Cache ID 6WYXi0OCq]
40. Chen Y, 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 [FREE Full text] [doi: 10.3310/hta16380] [Medline: 23046909]
41. Guerriero C, Cairns J, Roberts I, Rodgers A, Whittaker R, Free C. The cost-effectiveness of smoking cessation support delivered by mobile phone text messaging: Txt2stop. Eur J Health Econ 2013 Oct;14(5):789-797 [FREE Full text] [doi: 10.1007/s10198-012-0424-5] [Medline: 22961230]

Abbreviations

FTND: Fagerström Test for Nicotine Dependence

Edited by G Eysenbach; submitted 12.03.15; peer-reviewed by E Augustson; accepted 16.04.15; published 04.07.15.

Please cite as:
Borrelli B, Bartlett YK, Tooley E, Armitage CJ, Wearden A
Prevalence and Frequency of mHealth and eHealth Use Among US and UK Smokers and Differences by Motivation to Quit
J Med Internet Res 2015;17(7):e164
URL: http://www.jmir.org/2015/7/e164/
doi:10.2196/jmir.4420
PMID:26149323

©Belinda Borrelli, Yvonne Kiera Bartlett, Erin Tooley, Christopher J Armitage, Alison Wearden. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 04.07.2015. 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,

XSL•FO
RenderX

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

Trials of Intervention Principles: Evaluation Methods for Evolving Behavioral Intervention Technologies

David C Mohr[1], PhD; Stephen M Schueller[1], PhD; William T Riley[2], PhD; C Hendricks Brown[3], PhD; Pim Cuijpers[4], PhD; Naihua Duan[5], PhD; Mary J Kwasny[6], PhD; Colleen Stiles-Shields[1], MA, MS; Ken Cheung[7], PhD

[1]Center for Behavioral Intervention Technologies (CBITs), Northwestern University, Chicago, IL, United States

[2]Office of Behavioral and Social Sciences Research, National Institutes of Health, Washington DC, DC, United States

[3]Department of Psychiatry, Northwestern University, Chicago, IL, United States

[4]Department of Psychology and Pedagogy, Vrije Universiteit, Amsterdam, Netherlands

[5]Department of Psychiatry, Columbia University, New York, NY, United States

[6]Department of Preventive Medicine, Northwestern University, Chicago, IL, United States

[7]Mailman School of Public Health, Columbia University, New York, NY, United States

Corresponding Author:
David C Mohr, PhD
Center for Behavioral Intervention Technologies (CBITs)
Northwestern University
750 N. Lakeshore Drive, 10th Floor
Chicago, IL, 60611
United States
Phone: 1 312 503 1403
Fax: 1 312 503 1403
Email: d-mohr@northwestern.edu

Abstract

In recent years, there has been increasing discussion of the limitations of traditional randomized controlled trial (RCT) methodologies for the evaluation of eHealth and mHealth interventions, and in particular, the requirement that these interventions be locked down during evaluation. Locking down these interventions locks in defects and eliminates the opportunities for quality improvement and adaptation to the changing technological environment, often leading to validation of tools that are outdated by the time that trial results are published. Furthermore, because behavioral intervention technologies change frequently during real-world deployment, even if a tested intervention were deployed in the real world, its shelf life would be limited. We argue that RCTs will have greater scientific and public health value if they focus on the evaluation of intervention principles (rather than a specific locked-down version of the intervention), allowing for ongoing quality improvement modifications to the behavioral intervention technology based on the core intervention principles, while continuously improving the functionality and maintaining technological currency. This paper is an initial proposal of a framework and methodology for the conduct of trials of intervention principles (TIPs) aimed at minimizing the risks of in-trial changes to intervention technologies and maximizing the potential for knowledge acquisition. The focus on evaluation of intervention principles using clinical and usage outcomes has the potential to provide more generalizable and durable information than trials focused on a single intervention technology.

(J Med Internet Res 2015;17(7):e166) doi:10.2196/jmir.4391

KEYWORDS

mHealth; eHealth; clinical trials; methodology

Introduction

Background

Behavioral intervention technologies (BITs) employ technologies, such as mobile phones, tablets, computers, sensors, and other tools to support behavior change related to health, mental health, and wellness. Mobile health apps, treatment and prevention websites, sensors used in activity trackers, and smartwatches are common examples [1]. The term BIT is used, rather than eHealth or mHealth, as these terms can reflect a much broader area of medicine and informatics not necessarily focused on behavior change [2]. In practice, BITs change and evolve over time. As anyone who has installed an app knows,

XSL•FO
RenderX

their life on a device evolves through a steady stream of updates. Many of these updates are bug fixes, operating system upgrades, or changes to support the emergence of new devices. However, some changes alter the content and functionality and are intended to modify or improve the user's experience and the benefit they receive.

BITs change to harness affordances provided by the rapidly changing technological environment, such as improving computing power, leveraging new data capture and user interface functions, and growing capacity for data transmission [3]. The expectations and culture of BIT users are rapidly changing. When developers of a BIT observe that a feature used in other apps has become popular, they often add similar functionality to respond to the expectations of their customer base. For example, social networking and peer-to-peer messaging is a common feature in many recently developed BITs due to the popularity of social networking tools such as Facebook, Twitter, etc. Failing to meet changing user expectations relegates a BIT to increasing irrelevance to users. There have been increasing calls for methodologies that allow for continuous quality improvement through more rapid incorporation of changes and accumulating knowledge in the context of trials [4-6]. The purpose of this paper is to propose adaptations of traditional randomized controlled trial (RCT) methodology that can support evaluation of BITs.

The Purpose of a Randomized Controlled Trial

Modern RCT methodologies in medicine were developed to evaluate pharmacological agents that are not intended or expected to be modified frequently. To respond to early critiques that psychological interventions have little effect [7], rigorous methodology was developed, often using principles from pharmacological trial design, to evaluate psychological and behavioral interventions. Among the many innovations and adaptations, psychological treatments were standardized through manualization, therapist training, supervision, and fidelity monitoring to maximize internal validity [8]. These rigorous methods of "locking down" a psychological treatment contributed to a vast literature supporting the efficacy of psychological and behavioral interventions, as well as broad acceptance of their clinical value. However, it is doubtful that even with these methodologies, behavioral interventions are truly locked down to the degree a pharmacological agent is during a trial. And once moved to real-world clinical practice, they are rarely implemented in the same manner as in the RCT [9,10]. The application of RCT methodology developed for pharmacotherapies, translated for behavioral and psychological interventions, to BITs is even less appropriate.

There are two primary problems with locking down a BIT in an RCT. First, the length of time required to conduct an RCT is often not consistent with the rapidly changing technological environment [11] and user expectations. In some cases, trials of BITs can be conducted rapidly, however, many times trials must extend over longer periods of time. This is especially true when adequately powering a trial necessitates extended participant recruitment. The number of participants required to power a traditional RCT makes sense when the resulting intervention (if found successful) will have a long shelf life and benefit many users over time. In RCTs for BITs, however, this is rarely the case. RCTs often test early versions of BITs that often have to undergo revisions prior to implementation due to factors not under the developers' or researchers' control, such as changes in technological environment and contextual factors (eg, perceived attractiveness and usefulness of a BIT), which can impact use patterns or outcomes if not addressed [6]. Thus, locking down a BIT means that the information gained from a trial might not be useful for the implementation, even if this takes place shortly after the trial period.

Second, and perhaps more importantly, new information is gained within the trial that can be used for quality improvement [12]. Even with careful user-centered design and pilot testing, problems in the BIT design that impact the user's ability or willingness to perform tasks are often uncovered when it is deployed at scale for the full trial. As these deficiencies are discovered, locking down the BIT and refusing to improve it may undermine the chances of success, waste effort and resources, and compromise the relevance in the knowledge gained, if any.

Thus, BIT researchers are left with a difficult choice. They can (1) continue to investigate the locked-down version to maintain internal validity (with increasing understanding of the problems in the BIT and its growing irrelevance), (2) make changes to the BIT but not report them (reducing the reliability of the scientific literature and limiting the ability of other researchers to build on the work appropriately), or (3) make changes to the BIT and report them (raising the question "What is being tested?").

Current RCT methodology is based on the lock-down model, well suited to determining if a fixed intervention is more efficacious than a control condition. However, given the rate of technological advancement and changes in user expectations, we argue that testing a fixed BIT often has very little public health value because the shelf life of that fixed BIT is often relatively short. We argue instead for allowing modification to BITs if these changes are reported. The question of what is being tested can be addressed by specifying the principles (model or rules) underlying the BIT and conceptualizing the trial as testing those principles, and not the BIT itself.

Trials of Intervention Principles

We argue that trials of BITs should be viewed as experiments to test principles within that BIT that can then be more broadly applied by developers, designers, and researchers in the creation of BITs and the science behind technology-based behavioral intervention. As such, we refer to these trials as "Trials of Intervention Principles" (TIPs), as they test the theoretical concepts represented within the BIT, rather than the specific technological instantiation of the BIT itself. Similar logic has been applied in other instances of interventions to identify the elemental components contained within the interventions, to improve the ability to generalize across trials, and to guide intervention selection, creation, and evaluation [13-15]. We define a principle as a model or set of rules that defines how a group of behavioral strategies is instantiated in a BIT and how the use of that BIT leads to an outcome. For example, a trial might evaluate an app that aims to increase exercise by including

XSL•FO
RenderX

functionality that is intended to support goal setting, monitoring, and feedback. A TIP would allow for optimization of the BIT in the service of testing the principle. For example, the monitoring tools might prove less than ideal and thus evolve over time, moving from text-based free entry to selecting from a list, to passive sensing. Some of these evolutions might improve usability and thereby outcomes (eg, shifting from user-initiated entry to passive sensing could reduce user burden, thereby providing more feedback and increasing value to the user). On the other hand, improving usability could be detrimental, either because the improved usability in one part of the BIT creates problems in another (eg, shifting from user-initiated entry to passive sensing might not be able to sense certain activities, like bicycling, thus frustrating the user and limiting their subsequent use of the BIT), or because improved usability interferes with behavioral aims (eg, shifting from active behavioral logging to passive sensing may reduce agency or remove opportunities for the user to notice and learn about their behaviors). A TIP would allow for these iterations to improve the usability of the BIT based on use data, while continuing to test the principle that using a mobile phone app to promote goal setting, monitoring, and feedback to increase exercise is effective, based on the primary outcome data.

In light of this, TIPs test a set of intervention principles that define an underlying rule, model, or mechanism of action on which a BIT impacts its target outcome. The principles being tested may be based on clinical theory and/or methods of using technology to interact with users. TIPs require a methodology that provides guidance on operationalizing and analyzing these principles. Below we describe a framework for describing and operationalizing those principles, considerations related to making changes to a BIT, and evaluation methodologies.

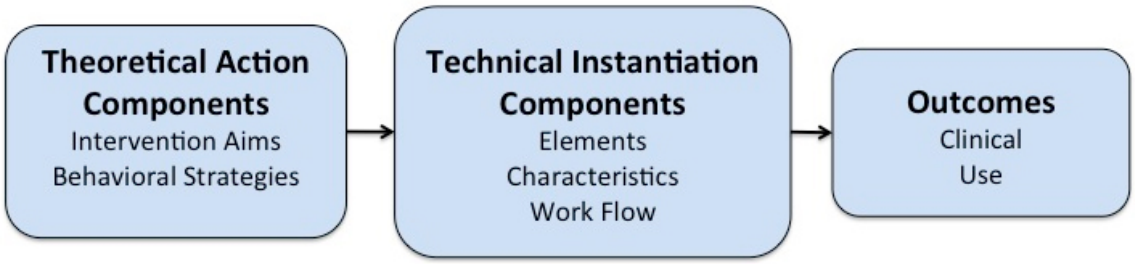
The Behavioral Intervention Technology Model

Overview

The principles to be tested must be defined and operationalized a priori to provide clarity to the aims, support change decisions during the trial (eg, defining what can and cannot be changed), as well as provide a method of documenting and reporting changes in a consistent manner. A framework for the definition of BIT principles can support investigators in characterizing these principles.

The BIT Model, shown in Figure 1, answers the questions “why”, “how”, “what”, and “when” of BIT development and includes two broad levels: (1) a theoretical action level, which reflects the intentions of the developer or researcher, and (2) an instantiation level, which reflects the technological implementation [16]. This model has been used to design and characterize intervention technologies [17].

Figure 1. The BIT Model.



Theoretical Action Components

The theoretical level includes two components: (1) intervention aims (note that in contrast to the BIT Model paper [16], the term “intervention” here refers to the entire treatment package), which are *why* the BIT exists, and (2) behavioral strategies, which are conceptually *how* the BIT will achieve those aims.

Intervention Aims

The intervention aims (“why” the BIT exists) reflect the fundamental intentions of the developer. These aims commonly include both explicit clinical aims (which may include preventive or well-being aims), such as weight reduction or reduction of depressive symptoms, as well as usage aims, which are sometimes implicit, such as expected frequency of use or the use of specific intervention elements. The clinical aims often have sub-aims. For example, a BIT that aims to reduce weight might aim to reduce caloric intake and increase physical activity. A BIT for depression may have sub-aims such as increasing positive activities and teaching cognitive restructuring. In general, the primary intervention aims, and frequently the sub-aims, are not changeable and are reflected in the assessment of primary outcome. Changing the primary aims of the BIT would require a new trial.

While clinical aims are most often mirrored by the primary outcome measurements, usage aims are more often expectations that the developer has regarding the level of use necessary for clinical effectiveness. While usage aims are not always highly correlated with outcome [18], they are often critical expectations, as they often provide early indicators of possible design flaws and may be used as indicators of the need for potential changes to the BIT.

Behavior Change Strategies

Behavior change strategies are the methods used to attain intervention aims and are grounded in models and theories of how behavior change occurs and is maintained. For example, Michie has developed a well-grounded taxonomy of such change strategies [19,20] such as education, goal setting, monitoring,

XSL•FO
RenderX

feedback, or motivational enhancement. Chorpita et al have similarly identified common practice elements among various behavioral and mental health interventions creating 26 codes in one instance and 47 in another [13,14].

Behavioral strategies are often key principles being tested in trials of apps (although this may not be true for trials with a human computer-interaction focus) and thus usually cannot be changed or eliminated during a trial. A decision to do so would prevent the investigator from being able to draw conclusions from the results of the trial.

Instantiation Components

Overview

The instantiation components include (1) BIT elements ("what" is delivered), (2) characteristics of those elements ("how" they are delivered), and (3) workflow ("when" they are delivered). In many cases these components can be tracked through use data. These three instantiation components are operationalized in the hardware specifications and software programming code of the BIT.

Elements

BIT elements reflect "what" is provided to the user. The elements are the distinct objects of a BIT intended to implement the behavior change strategies, which in turn support the user in achieving the clinical and usage aims. By BIT elements, we mean the actual technical instantiations present in the BIT. For example, a content delivery element might provide the user with information. Notifications are individual messages pushed to the user, such as text messages, emails, or within app notifications. Logging elements allow users to enter information. Reports and visualizations are reflections of data collected by the BIT that are provided back to the user (eg, calendars, calorie counts, thought records), often used to instantiate behavioral strategies such as feedback. Thus, the BIT elements are the aspects of the BIT with which the user actually interacts.

Some elements in a BIT may be part of the principles being tested, and others may not. For example, an investigator may be interested in testing one or more methods of logging tools for behavior, making these part of the principles being tested. Other elements may be included but are not of scientific interest to the study. For example, reminders may be included to support use but may not be part of the principles being tested. In this case, the investigator could not eliminate the logging elements but might be able to alter some of the characteristics or workflow within the parameters of the principles being evaluated, such as adding reminders to cue logging activities. However, the investigator would have much greater latitude in altering, adding, or eliminating the reminder elements, as long as the alterations did not introduce a new principle that might compete with the original principles.

Characteristics

Characteristics are "how" the elements are deployed. Elements can be considered objects, while characteristics are attributes of those objects. For example, an informational element may have a variety of characteristics, such as text, video, or audio content. That content may be simple or more complex. Form and esthetics are also characteristics of the elements. Elements may be personalized using characteristics to fit the demographics, language, or culture of the individual user.

In general, behavioral scientists have been less concerned with the characteristics, other than to make the BIT more usable and attractive to the user. However, in some instances the characteristics of BIT elements are the principle being investigated, for example, when evaluating the utility of personalized applications relative to static ones [21]. In those circumstances, substantive alterations to those characteristics under investigation may harm the trial.

Workflow

Workflow reflects "when" specific elements of a BIT are delivered. Most BITs are designed for repeated interactions over an extended period of time. The workflow determines when specific sets of elements are delivered, the sequence of the delivery, and the length of the intervention. User-defined workflows contain no coded rules, allowing the user access to all intervention elements and content. Thus, the user decides the sequence and timing of their use. Many BITs set conditions for when specific elements are made available. Time-based rules define the release of an intervention element based on the passage of time. For example, Web-based treatments modeled on standard face-to-face treatments sometimes release new content on a weekly basis [22,23]. Event-based rules define the release of elements based on the criteria detected by the intervention, such as the completion of a task or the detection of a user state. Just in time interventions may use complex workflows that use combinations of user characteristics, use data, detected events, and time to determine the delivery of intervention components [24].

Workflow as a principle has not received much attention in BIT research. However, workflow may impact principles that are being evaluated. For example, shifting notifications from fixed time-based to event-based may change the underlying behavior change strategy, as a notification triggered by behavior may be a form of feedback. Thus, changes to workflow as well as other theoretical and instantiation components may introduce factors that researchers did not anticipate.

Behavioral Intervention Technology Model in the Trials of Intervention Principles Framework

Overview

Figure 2 places the BIT Model into the context of TIPs. Aspects included within the dotted box are the main focus of the TIP and therefore should not be changed during the course of the trial. The box on the left, labeled "BIT", represents the instantiation features of the BIT, which define the intervention as deployed.

The central box represents the principles being tested in that deployment, which must always have an intervention aim and cannot change over the course of a trial. BITs designed by clinical or public health researchers almost always are based on some set of behavioral strategies, which often are also part of

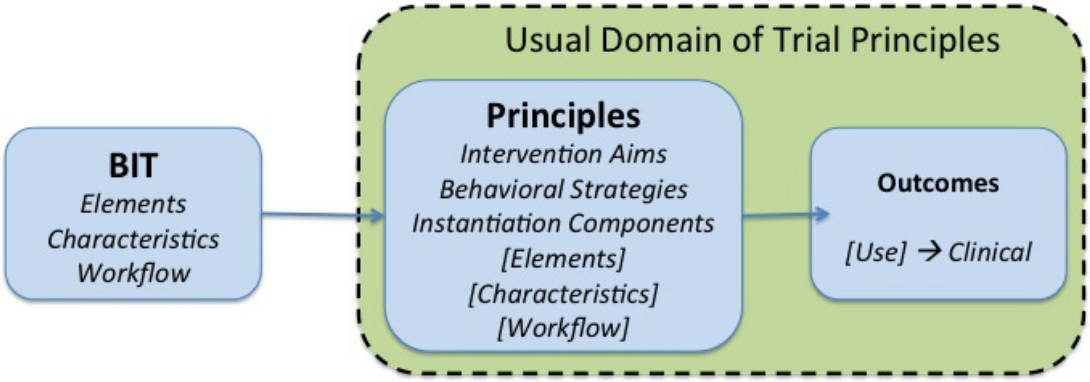
XSL•FO
RenderX

the intervention principles. Instantiation features (ie, elements, characteristics, or workflow) are represented in brackets because these may or may not be part of the principles of trials conducted by researchers, but often are the focus of trials with an engineering or human computer interaction focus. Instantiation components that are listed as principles may not be modified in the BIT during a trial. To the degree that those instantiation components are not part of the principles, or their alteration would not affect any of the principles, adaptation could be considered.

The outcomes box on the right represents the measurement of outcomes. The intervention outcomes generally also should not change. Use data (eg, use of a specific BIT element) can be monitored to indicate if the behavioral principles are being administered properly and are therefore often examined as more proximal measures of usability and the delivery of behavioral strategies [25]. Use data are often the criteria used to trigger considerations of changes to the BIT.

The TIPs framework can be viewed through the lens of mediational models in which an intervention is intended to affect intermediary outputs and outcomes [26]. The first part of the mediational model predicts that use of the BIT will increase the use, behavior, or experience defined by the principles. The second part of the mediational model predicts that the intervention principles will improve intervention outcomes. Thus, the principles and their measurement must remain fixed over the course of a trial. However, each new version of a BIT can be viewed as a moderator of this mediational model.

Figure 2. BIT Model in the context of TIPs.



Operationalizing Principles: The Principle Statement

The principles to be tested must be operationalized and stated a priori. This can be accomplished through a "Principle Statement" that describes each of the relevant BIT Model components [16] being evaluated, as shown in Figure 2. The formulation of a principle statement can be facilitated by using a model, such as the BIT Model, that integrates the conceptual and technological instantiation components. Using the BIT Model, a principle statement could take the following form: "The aim of the trial is to test an application that supports users in (*Behavioral Strategies*), delivered using (*BIT Elements, Characteristics, Workflow*) to improve (*Intervention Aim*)". A principle statement might not contain all of these aspects, as some BIT elements, characteristics, or workflow may not be considered principles in the trial. A principle statement is needed to provide clarity to the aims, support change decisions during the trial (eg, defining what can and cannot be changed), as well as provide a method of documenting and reporting changes in a consistent manner.

As an illustration, a principle statement for MyFitnessPal a few years ago might have read "MyFitnessPal aims to support users in goal setting, self-monitoring, feedback using logging, and data reporting and vizualization features to support *weight loss and increased physical activity*". This principle statement identifies the behavioral strategies and specifies a few BIT instantiation components linked to those strategies (ie, logging features for goal setting and monitoring, and data reporting and visualization to provide feedback on caloric intake and exercise), and measureable intervention aims (ie, weight loss and increased physical activity). In this statement, characteristics and workflow are not identified as principles being tested, thus, trialists would have been free to make changes to characteristics and workflow that did not substantially impact the behavioral strategies and BIT Model elements specified in the principle statement. MyFitnessPal's addition of the feature allowing the use of barcode scanning for data entry can be seen as consistent with the principle statement, as it is a characteristic that simplified data logging elements. The addition of social network features, however, added substantial new functionality. A critical determination should be made as to whether social network features would alter the trial principles of goal setting, self-monitoring, and feedback and whether they expand beyond the behavioral strategies of goal setting and self-monitoring. One could imagine social networking features framed around these topics. However, social networking also adds other social motivators such as accountability [27], solicitation, and receipt of advice. These would likely constitute the addition of a behavioral strategy, thereby undermining the interpretability of trial results with respect to the principles being tested.

The principles may be defined broadly, focusing primarily on intervention aims and behavioral strategies, or very narrowly, including BIT elements, characteristics, and workflow. The maximal amount of definition would come much closer to a trial of a locked down BIT, while a definition that included no BIT instantiation features would be indistinguishable from an intervention that did not include technology. Thus, the inclusion of BIT instantiation features in a principle statement should be

XSL•FO
RenderX

sufficient to define the intervention, but not include those areas that may allow for optimization in the service of testing the targeted principles.

Making Changes During a Trial of Intervention Principles

Mid-trial changes to a BIT will frequently be the result of knowledge gained up until that point [6]. Trialists and developers may use various sources of information to trigger considerations of whether changes to a BIT during trials should be made, including use data, user feedback, and possibly intervention outcomes. Use data can include gross measures of use that can provide general information on whether the BIT is being used, as well as detailed use data from specific BIT elements that can indicate how they are being used. Such data can be collected and reviewed periodically. User feedback is also commonly captured both through self-report surveys as well as through interviews with users at the end of treatment or at specified points during treatment [28].

The decision to change a BIT is almost by definition triggered by observations that were not anticipated prior to the launch of the trial. Thus, the decision is based most often on investigator judgment. The decision to make a change is often not easy, as investigators are often balancing pressure to succeed and budget constraints in a situation where there is often considerable uncertainty. While clear plans for such unforeseen situations would be difficult to develop, planning nonetheless may help facilitate decision making under these circumstances.

Usually a decision to change a BIT based on observed data is made due to a belief that a specific behavioral strategy, BIT element, characteristic, or workflow is producing suboptimal usage or intervention outcomes. Implicitly, this decision is based on a comparison between an a priori expectation and observation of usage or effect. However, in practice, a priori expectations are often not clearly defined, which in effect leaves investigators making decisions based on what they believe they would have expected beforehand, rather than actual prior expectations. Thus, the decision-making process can be facilitated by making the implicit more explicit—defining the expected use of specific BIT elements and outcomes prior to the app launch. These a priori expectations can be informed by pilot testing, prior experience, published data, or consultations with advisors. The decision to make a change may also be informed by examining the relationship of the use of the element with outcomes (intervention and usage), bearing in mind that some outcomes, especially more distal primary outcomes, may change more slowly than more proximal intermediary outcomes tied to specific behavior change strategies [25].

Degree of Blindness to Use and Outcome Data

Whether or not to keep investigators blind to some or all of the trial results during the trial and how to structure any lack of blinding are critical questions in managing TIPs. Investigator blindness is often thought to be a requirement for fixed interventions in order to protect against deliberate or unconscious biases. Full protection against this type of bias may be of less importance in quality improvement designs where there is a tacit if not explicit recognition that the BIT can be improved. There are situations where randomized trials explicitly use up-to-date outcome data to change certain aspects of a design (eg, adapting allocation ratios), but these must be used with caution to make appropriate causal inferences [29]. Also, it is possible to consider quality improvement changes in a BIT as discrete rather than continuous, with each instantiation being tested in one of a sequence of subtrials where blindness ends after each subtrial (see below under Evaluation). Alternatively, there are mechanisms for maintaining a degree of blindness throughout. In some cases, an investigator might have little involvement in the intervention being conducted and the eventual analysis of primary outcome data. In such cases, an unblinded BIT investigator, with blinded data collectors, analysts, and allocation concealment, has little potential to bias the research, short of engaging in research misconduct and might be the best person suited to review data to determine whether it is necessary to make changes and what those changes might be. Data from the control arms may be sequestered, thereby allowing investigators to monitor the experiences of BIT users but preventing them from knowing the results of the trial. This would allow for enough information to make decisions regarding changes but under the right controls could protect the evaluation from investigator bias.

Decision Making

Making in-trial changes to a BIT carries a number of risks and should be undertaken only when deficiencies are observed. The goals of making changes are to improve the BIT's capacity to achieve the primary intervention goals of the intervention using the intervention principles being tested, to prevent obsolescence of the BIT that may occur in the changing technological and application environment, and to maximize the amount of knowledge obtained from the substantial costs and efforts put into a trial. Bug fixes are probably the least controversial types of changes that can and are made during trials. Bug fixes are problems in the functionality of the BIT that might arise from changes in the technological environment (eg, changes in operating systems that cause problems for an app) and problems encountered in wider deployment not uncovered during initial testing. Other changes, however, implemented to improve functionality and usability open additional risks that can undermine TIPs. These risks can occur in at least three ways.

1. The proposed changes may directly diminish the BIT's ability to provide the intervention principles by reducing its usability or usefulness. While investigators and developers would not likely consider such deleterious changes, this may occur inadvertently. The example noted above (the addition of reminder notifications) may increase initial response to prompts and engagement; however, this may also engage people who do not sustain engagement, thereby decreasing overall retention [30]. Therefore, researchers should carefully consider unintended consequences of a supposed improvement and intensively monitor potential changes resulting from this improvement to insure that it has improved the intervention.
2. Changes to a BIT may enhance components of the app that provide an alternative pathway to the intervention aims that undermines the primary principles being tested. For example, adding a motivational messaging element to an

XSL•FO
RenderX

app being used to test the value of goal setting and feedback could introduce another behavioral strategy that would undermine the trial's interpretability. Furthermore, the impact of changes may not result from a single decision but may be the consequence of a series of incremental changes that each by itself do no harm, but which in their totality undermine the trial.

3. A change might increase variability in how people use and benefit from a BIT. For example, in a BIT to increase physical activity through passive sensing and monitoring, developers might introduce a leaderboard that provides feedback about how one's steps compares to other similar users. Leaderboards tend to be a polarizing feature, with some people motivated by such competition and therefore increasing the target behavior, while other people are negatively impacted (eg, experiencing negative emotions or loss of intrinsic motivation) thus decreasing use of the BIT and the target behavior [31]. In such a case, the BIT could still have the same impact for the "average" user, but as users either increase or decrease their use and activity in response to the feature, this would increase the variability of response. Increased variability in outcome could make it more likely that the null hypothesis (of no difference between groups) would be supported.

To prevent such problems, a clear set of procedures should be employed in making such decisions. First, the investigators should consider how the changes would affect the interpretation of the results of the trial. Most importantly, they should consider whether or not the proposed changes would open alternative explanations to any trial results. Second, any proposed changes should not just be considered in isolation but in the totality of all changes made to date. That is, the proposed changes would impact the interpretability of the results relative not only to the previous iteration, but also to the first deployed version of the BIT.

A number of questions can be considered when weighing the pros and cons of making an in-trial change to a BIT:

- Does the change interfere with a primary principle being tested?
- Would the proposed change alter the principle statement?
- What is the operational definition of success for a proposed change? Often the definition of success can be defined by the usage aims, which can be observed in the short-to-intermediate term. However, success may also be defined with respect to the intervention aims, which are more relevant but frequently require longer follow-up.
- If the introduction of the change is successful, would it create an alternative explanation for the success of the trial?
- If an alternative explanation is possible, are there methods that can be used that can reasonably eliminate that alternative explanation?
- If a change is not made, how would this impact the generalizability/implementation of findings from the trial?

In making a decision to change a BIT during a trial, input from a variety of perspectives is important to ensure that all possible impacts are considered. The areas of domain expertise brought to bear on the question may include behavioral science and theory, clinical or public health expertise, human factors engineering, trial methodology, statistics, and any other relevant area. It may also be useful to have guidance from knowledgeable people who are not part of the core investigative team, who may be able to see potential problems that a team, involved in the day-to-day operations, is less likely to see. If a panel of stakeholders, including potential users, were created to provide feedback during BIT design, this panel could be consulted on changes considered during the trial as well. Some decision-making functions could be embedded in the data safety monitoring board, if this board can operate in an efficient manner.

Once possible threats of a proposed change have been identified, methods of mitigating those threats to interpretability can be considered. For instance, using our previous example, if motivational messaging is added to an app with the intention of improving use, the impact of such a decision on a trial might be mitigated by adding identical or similar messaging to a BIT used in a control condition. Furthermore, subsequent investigations might be useful to demonstrate that a proposed principle is generalizable and persists when aspects not related to that principle are modified.

Documentation and Reporting

Documentation is critical to TIPs of evolving technologies both to support the investigative team in decision making and in creating transparency. The principles being tested should be defined before the start of the trial, for example, using a principle statement. In general, the greater the specificity of the principle statement, the less latitude the investigator and developer will have in making changes.

Finally, documentation of the changes and the reasoning behind them will provide transparency and allow consumers of trial data to understand and accurately interpret the meaning of the trial results. While providing an accurate definition of the intervention in any peer-reviewed publication is important [32], reporting should also include changes made to BIT elements, characteristics, and workflow (and behavioral strategies if necessary) over the course of the trial, in accordance with the CONSORT-EHEALTH Guidelines [33].

We would like to emphasize, that in most cases, changes to BITs are relatively minor and infrequent. The cost of making continual or substantial changes to BITs is often a significant limiting factor. In addition, collecting, collating, and reviewing data that would drive these changes is time intensive and will also limit the frequency of such changes. Even without these barriers, researchers should be careful not to "churn" or overcorrect when making changes to the intervention elements, characteristics, or workflows. As easily implemented improvements are generated, these should be considered carefully, grouped into meaningful sets, and implemented as a set with appropriate version control of the software.

Evaluation

Overview

The evaluation of RCT data on the efficacy or effectiveness of BIT in its entirety as an application can be performed using

XSL•FO
RenderX

standard statistical analyses such as paired *t* tests or more complicated analytical procedures such as mixed effects modeling. BIT use can be reported using simple descriptive statistics such as the mean or media of app launches or site logins, cleaned for artifacts, and between-application comparison can also be performed in an unbiased fashion using comparative tests such as *t* tests, Mann-Whitney U tests, or mixed effects modeling. The advantage of evaluating the efficacy of a BIT as an application is its simplicity and usefulness for between-BIT comparisons. However, a standard analytic approach ignoring changes made to a BIT essentially aggregates the effects of those changes over time, combining the effects of the first version with the results of each of the subsequent iterations.

Verification of Improvements

An important part of any quality improvement effort is the verification that the changes made to the BIT had the intended effects. This verification can be performed both on the intervention outcomes and the use data. In most cases, such analyses are more likely to be better powered for use data than for primary outcomes, as use data are usually more proximal to the changes in BITs intended to improve usability. This can be evaluated by comparing use data (or outcomes) across different versions of the app or by testing the fixed effect of a time-varying ordinal component denoting the version in a longitudinal design. Such evaluations would need to check for changes in the sample (violations of the assumption of stationarity) that might occur over time. If changes in the recruited sample occurred over time, these changes would need to be addressed using covariates in the analysis or correcting through a propensity score analysis [34].

Optimization Trials for Local Evaluation

In practice, when a team has reasonable certainty on an approach to take, a specific change in the technology can be made, and the effect of that change can be monitored, as described above. However, it is not uncommon that development teams have more than one viable change and do not have enough information to make a decision. In such cases, optimization trials may be conducted within the treatment arm. The goal of an optimization trial is to maximize the outcomes of the trial participants, strengthen the test of the principles being evaluated, and support learning during the trial. A simple approach would be to randomize participants to versions of the BIT that contain the different solutions until the developers are confident that one of the solutions is superior.

Adequate power to detect differences would in many cases require large effect sizes under a traditional accuracy-centric framework that controls the study accuracy via type I error rate. This would normally not be possible for clinical outcomes, where sample sizes are determined by power analyses for the entire trial and may be a challenge even for use data (although those effect sizes should be larger if one of the options being tested is clearly superior). However, if optimization of a BIT is considered a local evaluation, within the trial, to support maximization of BIT use and/or outcomes of the trial participants, local trials and a selection paradigm may be more appropriate [35]. In this framework, assuming prior knowledge does not indicate a clear preference. A decision critical value may be defined with respect to a 50% type I error rate (ie, requiring P value <.50), thus imposing substantially smaller sample sizes. The one caveat is that the knowledge acquired from the decisions are intended for local optimization and cannot be generalized beyond the specific test of the application and should be used only for validating the effects of optimization efforts within the deployment for that trial.

Trials Involving Individual Level Adaptation Strategies

An important characteristic of BITs is that they can learn and respond to individual use patterns and self-reports or sensors of certain health outcomes. One could adopt outcome-adaptive randomization that uses the data from patients previously treated in the trial to tilt the randomization probabilities in favor of the instantiation components (elements, characteristics, and workflow) having comparatively superior outcomes such as maximum usage [36]. This adaptive randomization provides a balance between ethical concerns and the learning objective. Importantly, the adaptive randomization probabilities can be personalized based on a patient's covariate or profile [37]. While statistical inferences (eg, hypothesis testing) after an outcome-adaptive procedure should be interpreted with caution, adaptive randomization is useful for optimizing in-trial outcomes and for selection purposes. In order for the adaptation to come into effect, we need to assign initial patients in a non-adaptive fashion (such as balanced randomization) so that we use the outcome data in these initial patients to change the randomization probabilities. However, the expected outcomes of patients will improve over time as the trial continues, and accrual and adaptation begin. In addition, in order to shorten the wait period before adaptation begins, adaptation can be made on the basis of early outcome or usage data that are indicative of the eventual outcome [38].

In the context of TIPs, randomization and treatment assignments can be made multiple times within each patient as in a crossover study design. Furthermore, the reassignment can be made based on the intermediate outcome or usage data as in a sequential multiple assignment randomized trial (SMART) design [39,40]. For example, suppose a patient is first assigned to receive no reminder and has not used the BIT for 2 weeks. Then the patient is reassigned to receive daily reminders and begins to check in to the app in response to the reminder. Then the frequency may be reduced to three times a week so as to avoid "overdose". Generally, the data structure from a SMART design is dynamic in that the interventions are dependent on the intermediate response. For this type of data, we could use reinforcement learning techniques such as Q-learning to identify the optimal dynamic sequence of actions [41]. Q-learning involves fitting a regression model for a proxy of the eventual outcome at each decision stage so as to assess the impact of change and its interaction with the patient's history (such as profile and intermediate response), and using backward induction to obtain the stage-wise optimal decisions. By virtue of randomization in a SMART, such analysis provides unbiased evaluation of the components and principles within a BIT. Importantly, on the basis of the evaluation, we could assess the optimal implementation of a BIT using backward induction.

XSL•FO
RenderX

To evaluate the effects of individual intervention components in complex interventions such as a BIT, n-of-1 trials can also be useful [42], in particular when the effects are stationary over time but heterogeneous across patients. While both n-of-1 trials and SMART involve re-randomization within a patient, their objectives are different. In a SMART, the goal is to identify a *sequence* of intervention components so as to maximize an eventual health outcome; in an n-of-1 trial, the goal is to identify an *intervention component* that is effective in a particular patient. These two designs provide complementary tools in the evaluation of BITs.

Evaluating Principles

The strongest test of a principle is to withhold it from some subjects and evaluate whether there is a difference in outcomes between those who received that principle and those who did not. Most BITs, however, rely on multiple behavioral strategies. One can evaluate the behavioral strategies as a package, much as is commonly done in trials of behavioral and psychological interventions. Alternatively, trials using multiphase optimization strategies and factorial designs can be used to test a number of principles in a single trial [43].

For trials evaluating multiple principles as a package, post hoc evaluations may be performed to test the relative contribution of different components to outcomes. For example, if an app relies on goal setting, monitoring, and feedback as change principles, one could evaluate the relative use of each of the components to outcome. Such designs are purely correlational and are therefore not conclusive, but they may provide information that could be valuable to future investigations.

Development of Statistical Methodology for Trials of Intervention Principles

While many statistical methods are available that can be used for the evaluation of TIPs, approaches that truly integrate optimization methods and RCT analysis will be required. Methodology used for cumulative trials may be useful in this context. A "cumulative trial" is a sequence of smaller trials testing single or multiple adaptations in the technology that can be analyzed in a combined fashion relative to a control condition [29,44]. In the context of TIPs, a cumulative trial could involve a sequence of discrete quality improvement tests, intended to improve BIT performance, compared against a single control condition that can test underlying principles. This proposed method evaluates the effect of the BIT on the intervention outcome, the comparative effect of individual BIT versions, and the effects of varying levels of use of the BIT on outcomes across and within BIT versions.

As an example of a potential cumulative TIP design, we consider how such a design might have been used during the early period of MyFitnessPal. This proposed TIP design would model the underlying intervention impact of the principle as a function of the effectiveness of the BIT to affect outcomes. In symbols, we consider a sequence of trials, t=1 , …, T, each of which is testing a control condition against one or more versions of active BITs. For simplicity, suppose there is a single active version that is modified in each new trial t. Each individual i in trial t is randomly assigned to either active intervention, where Z_{it}=1 or control Z_{it}=0. For each of the trials, we suppose there is a common use metric, that is, U_{it},=1 if person i achieves full usage, 0 if no usage, and between these extremes for partial usage. In this case, usage would include the degree of reporting of meals and exercise. Let Y_{it} be the measured health outcome for subject i in trial t; in this case it would be weight loss. The effect of the active version compared against control could be modeled as involving three components in a multiplicative model. The first component involves a factor whose strength is measured by the parameter β that compares the BIT versus control under a condition where the subject receives the optimal version and the individual has complete usage. The second component's strength is measured by a parameter γ_t for all other versions that accounts for their being suboptimal. This component, given within the first brackets below, is bounded between zero and one. The third component, parameterized by δ, accounts for degradation in effects due to incomplete usage by the individual. BIT use can be calibrated on a zero to one scale using simple descriptive statistics such as the mean or median of app launches or site logins, cleaned for artifacts, and controls would receive a zero value unless they self-expose to an alternative BIT. This component is the third bracketed expression below, and it too is bounded between zero and one:

$$Y_{it}=\mu + Z_{it}\ \beta * [\exp(\gamma_t) / 1+ \exp(\gamma_t)] * [\exp(U_{it}\ \delta) / 1+ \exp(U_{it}\ \delta)] + \varepsilon_{it} \quad (1)$$

The last term in equation (1), ε_{it}, corresponds to an error term. We can interpret β as the optimal effect of a BIT against control on the intervention outcome; the γ_t represents how close to optimal this version is, and δ can be interpreted as the sensitivity that usage or dosage has on the health outcome. The degree of usage itself can be modeled as well, such as:

$$\log (U_{it} / 1 + U_{it})= \theta + \upsilon_t + \varepsilon'_{it} \quad (2)$$

In this modeling equation (2), the value of υ_t represents the average usage on this transformed scaled for the t^{th} version. These models are illustrations of how suboptimality and usage can be included and would likely need refining so that they provide a good fit to the data from a sequence of trials. Statistical inference across trials would need to account for trial level variation, much like meta-analysis, network meta-analysis, and other synthesis methods include trial level random effects [44]. In addition, individual level factors would need to be included both as main effects and potential moderator variables. Designs that deliberately look at variation in impact as a function of baseline participant characteristics are useful for examining this moderation of effects on both use and intervention outcomes. To ensure sufficient power to detect moderation, participants are often stratified by baseline characteristics (eg, age) and then randomized within these strata.

Equations 1 and 2 can be used to conduct a type of mediation analysis that partitions how much of the impact is explained through usage. Importantly, random effects can be entered into equation (1) to account for imperfect specification of a hypothesized mediation pathway. As this equation stands without random effects, all the effects of the intervention are required to flow through a specific mediation model (eg, usage).

XSL•FO
RenderX

Incorporation of random effects provides the opportunity to address more complex explanations of an intervention's effect. Also, classic mediation modeling [26] can be used to test whether an intervention's impact on clinical outcomes is due to the use of the principles underlying the BIT. Such mediation analyses are most useful when two versions of a BIT are tested against each other. A mediation analysis in a head-to-head trial of two BITs would first test whether use of the BIT varies, then test whether variation in use affects the more distal health outcome.

Discussion

Principal Considerations

Developers frequently make changes and adjustments to BITs during the course of trials for numerous reasons. Often original specifications are created with insufficient information. Even after careful usability and field testing, problems can be uncovered during broader deployment in a trial. Locking down an intervention during a trial, thereby forcing investigators to test a BIT with known deficiencies, decreases the likelihood that useful knowledge will come from the trial. Furthermore, the resulting BIT may not be appropriate for the technological environment that exists when the trial is completed. Thus, integrating quality improvement into RCTs can spur innovation, allow for BITs to evolve over time, and improve the likelihood that useful information will result from the trial.

We propose that these trials be conceptualized as evaluations of principles rather than of BITs themselves. As BITs, unlike pharmaceutical agents, do not stay static over time, the principles being evaluated represent the more static and generalizable component of the trial. This is valuable as the current form of any BIT at any given time will likely be relatively short lived. These trials, which we call TIPs, would evaluate constructs that could be more broadly applied, thereby moving the field ahead more rapidly.

We have presented an initial proposal of a method of defining the principles being evaluated that incorporates the essential clinical features, including the clinical aims and behavioral strategies, and the technical features, including BIT elements, characteristics, and workflow. This principle statement defines what is being tested, thereby promoting clarity for the research team and communication of aims to consumers of trial information. The principle statement can also support investigators in decision making around in-trial changes to the BIT. Changes in seemingly unrelated elements, however, may still have secondary effects on the principles under evaluation. Careful consideration of these potential secondary impacts is required to preserve the interpretability of the results. We recognize that even with these procedures, the risk of damaging a trial through in-trial changes to a BIT cannot be completely eliminated. However, we argue that the potential risks are far less than the more certain risks of persisting in a trial with a BIT that is underperforming in critical areas.

This approach fits well with changes in many funding agencies where there is a growing emphasis on trials that focus on mechanisms that can be more broadly applied [45]. Industry often has needs that are different from academic researchers, in that they are seeking to acquire data to support the effectiveness of their BITs for marketing purposes. But there too, the BIT that is purchased at one point is likely to change over time. A clear definition of the principles tested would allow the purchasers of BITs to have a clear definition of the parameters of the BIT they are purchasing, thereby allowing them to determine when the BIT they are receiving no longer meets the specifications of the BIT they purchased.

This is by no means the first proposal to enhance the usual RCT design to address the needs of BIT evaluation. Multiphase Optimization Strategy trials (MOST) have become an accepted method of using fractional factorial designs to evaluate and optimize the components of BITs [39]. SMART designs allow for the evaluation of systems that provide different components at various times depending on specified criteria [39,40] and thus have considerable value in investigating questions such as workflow. Continuous evaluation of evolving behavioral intervention technologies (CEEBIT) is a proposed method of comparatively evaluating BITs in local settings, a form of post-marketing surveillance allowing organizations providing BITs to continuously evaluate their usefulness in among their consumers [5]. TIPs extend this process of RCT innovation by providing a methodology for harnessing knowledge of BITs gained during the trial and managing changes in BIT specifications. TIP methodology is intended to prevent such changes from damaging the interpretability of results, while maximizing the amount of knowledge gained. Additional work might be undertaken after a principled trial to ensure the generalizability of the principles tested. However, these steps should be seen after initial trials provide sufficient evidence that these principles are being closely tied to behavior change.

TIP methodology is not intended as a substitute for careful usability and field testing prior to a trial. Ensuring that a BIT is usable, both through laboratory and field testing, is an essential step in acquiring information, refining the BIT, and ensuring that the tools are usable and useful for their intended consumers. Furthermore, information gained from laboratory testing cannot be obtained easily in field deployment. Early field testing can uncover fundamental problems that reduce the effectiveness of the BIT in promoting the intended behavioral strategies. Indeed, initiating a trial with a poorly developed app greatly increases the risk that a trial will fail, since early users will be less likely to show benefit, which will decrease the trial's power to see changes in the primary outcomes. Additionally, adopting TIPs in practice is not without its challenges. We have used the BIT Model [16] to identify levels and elements contained within a BIT and to emphasize the common behavioral strategies, the conceptual "how" of a BIT, used to attain the intervention goals, but other models could also be applied within the TIP framework. The underlying notion, however, that intervention trials are truly testing common elements rather than specific instantiations, has been emphasized with other models [13-15,19,20], and we have highlighted considerations on how to make use of these conceptual notions within trials of BITs. We believe adoption of TIPs is an initial step to increase the transparency of the evaluation process for technology-based interventions and improve the value of the information gleaned

XSL•FO
RenderX

from these trials. While further work is needed, this initial framework provides a common language and practices for the field to consider.

Conclusion

To the best of our knowledge, this is the first methodology proposed to manage changes to a BIT during the course of a trial. The proposed methods are intended to ensure that the principles under evaluation are preserved and even enhanced by allowing learning and optimization to occur during the trial.

Acknowledgments

We are grateful for support from the United States National Institutes of Health, including P20 MH090318, R01 MH095753, R01 MH100482, and R34 MH095907 (PI Mohr); P30DA027828 and R01MH040859 (PI Brown); R01 NS072127 (PI Cheung); K08 MH102336 (PI Schueller); F31 MH106321 (PI Stiles-Shields). The content of this paper is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

Conflicts of Interest

None declared.

References

1. Mohr DC, Burns MN, Schueller SM, Clarke G, Klinkman M. Behavioral intervention technologies: evidence review and recommendations for future research in mental health. Gen Hosp Psychiatry 2013;35(4):332-338 [FREE Full text] [doi: 10.1016/j.genhosppsych.2013.03.008] [Medline: 23664503]
2. 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]
3. Christensen C, Johnson C, Horn M. Disrupting class: how disruptive innovation will change the way the world learns. New York: McGraw-Hill; 2008.
4. Kumar S, Nilsen WJ, Abernethy A, Atienza A, Patrick K, Pavel M, et al. Mobile health technology evaluation: the mHealth evidence workshop. Am J Prev Med 2013 Aug;45(2):228-236 [FREE Full text] [doi: 10.1016/j.amepre.2013.03.017] [Medline: 23867031]
5. Mohr DC, Cheung K, Schueller SM, Hendricks BC, Duan N. Continuous evaluation of evolving behavioral intervention technologies. Am J Prev Med 2013 Oct;45(4):517-523 [FREE Full text] [doi: 10.1016/j.amepre.2013.06.006] [Medline: 24050429]
6. Riley WT, Glasgow RE, Etheredge L, Abernethy AP. Rapid, responsive, relevant (R3) research: a call for a rapid learning health research enterprise. Clin Transl Med 2013;2(1):10 [FREE Full text] [doi: 10.1186/2001-1326-2-10] [Medline: 23663660]
7. Eysenck HJ. The effects of psychotherapy: an evaluation. J Consult Psychol 1952 Oct;16(5):319-324. [Medline: 13000035]
8. Bellg AJ, Borrelli B, Resnick B, Hecht J, Minicucci DS, Ory M, Treatment Fidelity Workgroup of the NIH Behavior Change Consortium. Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. Health Psychol 2004 Sep;23(5):443-451. [doi: 10.1037/0278-6133.23.5.443] [Medline: 15367063]
9. Glasgow RE, Davidson KW, Dobkin PL, Ockene J, Spring B. Practical behavioral trials to advance evidence-based behavioral medicine. Ann Behav Med 2006 Feb;31(1):5-13. [doi: 10.1207/s15324796abm3101_3] [Medline: 16472033]
10. Chambers DA, Glasgow RE, Stange KC. The dynamic sustainability framework: addressing the paradox of sustainment amid ongoing change. Implement Sci 2013;8:117 [FREE Full text] [doi: 10.1186/1748-5908-8-117] [Medline: 24088228]
11. Hilbert M, López P. The world's technological capacity to store, communicate, and compute information. Science 2011 Apr 1;332(6025):60-65 [FREE Full text] [doi: 10.1126/science.1200970] [Medline: 21310967]
12. Baker TB, Gustafson DH, Shah D. How can research keep up with eHealth? Ten strategies for increasing the timeliness and usefulness of eHealth research. J Med Internet Res 2014;16(2):e36 [FREE Full text] [doi: 10.2196/jmir.2925] [Medline: 24554442]
13. Chorpita BF, Daleiden EL. Mapping evidence-based treatments for children and adolescents: application of the distillation and matching model to 615 treatments from 322 randomized trials. J Consult Clin Psychol 2009 Jun;77(3):566-579. [doi: 10.1037/a0014565] [Medline: 19485596]
14. Chorpita BF, Daleiden EL, Weisz JR. Identifying and selecting the common elements of evidence based interventions: a distillation and matching model. Ment Health Serv Res 2005 Mar;7(1):5-20. [Medline: 15832690]
15. Embry DD, Biglan A. Evidence-based kernels: fundamental units of behavioral influence. Clin Child Fam Psychol Rev 2008 Sep;11(3):75-113 [FREE Full text] [doi: 10.1007/s10567-008-0036-x] [Medline: 18712600]

XSL•FO
RenderX

16. Mohr DC, Schueller SM, Montague E, Burns MN, Rashidi P. The behavioral intervention technology model: an integrated conceptual and technological framework for eHealth and mHealth interventions. J Med Internet Res 2014;16(6):e146 [FREE Full text] [doi: 10.2196/jmir.3077] [Medline: 24905070]
17. Anderson JK, Wallace LM, Clark A, Regent L. Applying the Behavioural Intervention Technologies model to the development of a smartphone application (app) supporting young peoples' adherence to anaphylaxis action plan. BMJ Innovations 2015 Apr 02;1(2):67-73. [doi: 10.1136/bmjinnov-2014-000016]
18. Warmerdam L, van SA, Twisk J, Riper H, Cuijpers P. Internet-based treatment for adults with depressive symptoms: randomized controlled trial. J Med Internet Res 2008;10(4):e44 [FREE Full text] [doi: 10.2196/jmir.1094] [Medline: 19033149]
19. Michie S, Abraham C, Eccles MP, Francis JJ, Hardeman W, Johnston M. Strengthening evaluation and implementation by specifying components of behaviour change interventions: a study protocol. Implement Sci 2011;6:10 [FREE Full text] [doi: 10.1186/1748-5908-6-10] [Medline: 21299860]
20. 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]
21. Resnicow K, Davis RE, Zhang G, Konkel J, Strecher VJ, Shaikh AR, et al. Tailoring a fruit and vegetable intervention on novel motivational constructs: results of a randomized study. Ann Behav Med 2008 Apr;35(2):159-169. [doi: 10.1007/s12160-008-9028-9] [Medline: 18401673]
22. Mohr DC, Duffecy J, Ho J, Kwasny M, Cai X, Burns MN, et al. A randomized controlled trial evaluating a manualized TeleCoaching protocol for improving adherence to a web-based intervention for the treatment of depression. PLoS One 2013;8(8):e70086 [FREE Full text] [doi: 10.1371/journal.pone.0070086] [Medline: 23990896]
23. Andersson G, Bergström J, Holländare F, Carlbring P, Kaldo V, Ekselius L. Internet-based self-help for depression: randomised controlled trial. Br J Psychiatry 2005 Nov;187:456-461 [FREE Full text] [doi: 10.1192/bjp.187.5.456] [Medline: 16260822]
24. Brown CH, Mohr DC, Gallo CG, Mader C, Palinkas L, Wingood G, et al. A computational future for preventing HIV in minority communities: how advanced technology can improve implementation of effective programs. J Acquir Immune Defic Syndr 2013 Jun 1;63 Suppl 1:S72-S84 [FREE Full text] [doi: 10.1097/QAI.0b013e31829372bd] [Medline: 23673892]
25. Klasnja P, Consolvo S, Pratt W. How to evaluate technologies for health behavior change in HCI research. 2011 May 7-12 Presented at: SIGCHI Conference on Human Factors in Computing Systems (CHI); May 2011; Vancouver, Canada p. 7-12 URL: http://dl.acm.org/citation.cfm?id=1979396 [doi: 10.1145/1978942.1979396]
26. MacKinnon D. Introduction to statistical mediation analysis. New York: Lawrence Erlbaum Associates/Taylor & Francis Group; 2008.
27. Mohr DC, Cuijpers P, Lehman K. Supportive accountability: a model for providing human support to enhance adherence to eHealth interventions. J Med Internet Res 2011;13(1):e30 [FREE Full text] [doi: 10.2196/jmir.1602] [Medline: 21393123]
28. Ben-Zeev D, Schueller SM, Begale M, Duffecy J, Kane JM, Mohr DC. Strategies for mHealth research: lessons from 3 mobile intervention studies. Adm Policy Ment Health 2015 Mar;42(2):157-167. [doi: 10.1007/s10488-014-0556-2] [Medline: 24824311]
29. Brown CH, Ten Have TR, Jo B, Dagne G, Wyman PA, Muthén B, et al. Adaptive designs for randomized trials in public health. Annu Rev Public Health 2009;30:1-25 [FREE Full text] [doi: 10.1146/annurev.publhealth.031308.100223] [Medline: 19296774]
30. Goritz AS, Crutzen R. Reminders in Web-Based Data Collection: Increasing Response at the Price of Retention? American Journal of Evaluation 2011 Sep 20;33(2):240-250. [doi: 10.1177/1098214011421956]
31. Boyce A, Doran K, Campbell A, Pickford S, Culler D, Barnes T. BeadLoom Game: adding competitive, user generated, and social features to increase motivation. 2011 Jun 29 Presented at: 6th International Conference on Foundations of Digital Games; June 29-July 1, 2001; Bordeaux, France URL: http://dl.acm.org/citation.cfm?id=2159384
32. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014;348:g1687 [FREE Full text] [Medline: 24609605]
33. Eysenbach G, CONSORT- E. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
34. Harder VS, Stuart EA, Anthony JC. Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. Psychol Methods 2010 Sep;15(3):234-249 [FREE Full text] [doi: 10.1037/a0019623] [Medline: 20822250]
35. Cheung K, Duan N. Design of implementation studies for quality improvement programs: an effectiveness-cost-effectiveness framework. Am J Public Health 2014 Jan;104(1):e23-e30 [FREE Full text] [doi: 10.2105/AJPH.2013.301579] [Medline: 24228672]
36. Berry DA, Eick SG. Adaptive assignment versus balanced randomization in clinical trials: a decision analysis. Stat Med 1995 Feb 15;14(3):231-246. [Medline: 7724909]

XSL•FO
RenderX

37. Cheung Y, Inoue Lurdes YT, Wathen JK, Thall PF. Continuous Bayesian adaptive randomization based on event times with covariates. Stat Med 2006 Jan 15;25(1):55-70. [doi: 10.1002/sim.2247] [Medline: 16025549]
38. Cheung K, Kaufmann P. Efficiency perspectives on adaptive designs in stroke clinical trials. Stroke 2011 Oct;42(10):2990-2994 [FREE Full text] [doi: 10.1161/STROKEAHA.111.620765] [Medline: 21885845]
39. Collins LM, Murphy SA, Strecher V. The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): new methods for more potent eHealth interventions. Am J Prev Med 2007 May;32(5 Suppl):S112-S118 [FREE Full text] [doi: 10.1016/j.amepre.2007.01.022] [Medline: 17466815]
40. Lavori PW, Dawson R. Dynamic treatment regimes: practical design considerations. Clin Trials 2004 Feb;1(1):9-20. [Medline: 16281458]
41. Murphy SA. A Generalization Error for Q-Learning. J Mach Learn Res 2005 Jul;6:1073-1097 [FREE Full text] [Medline: 16763665]
42. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. Int J Nurs Stud 2013 May;50(5):587-592. [doi: 10.1016/j.ijnurstu.2012.09.010] [Medline: 23159157]
43. Collins LM, Murphy SA, Nair VN, Strecher VJ. A strategy for optimizing and evaluating behavioral interventions. Ann Behav Med 2005 Aug;30(1):65-73. [doi: 10.1207/s15324796abm3001_8] [Medline: 16097907]
44. Brown CH, Sloboda Z, Faggiano F, Teasdale B, Keller F, Burkhart G, et al. Methods for synthesizing findings on moderation effects across multiple randomized trials. Prev Sci 2013 Apr;14(2):144-156 [FREE Full text] [doi: 10.1007/s11121-011-0207-8] [Medline: 21360061]
45. Insel T. Director's blog: a new approach to clinical trials. 2014. URL: http://www.nimh.nih.gov/about/director/2014/a-new-approach-to-clinical-trials.shtml [accessed 2015-02-26] [WebCite Cache ID 6WdIwFEVG]

Abbreviations

BIT: behavioral intervention technology
RCT: randomized controlled trial
SMART: sequential multiple assignment randomized trial
TIP: trial of intervention principles

Edited by G Eysenbach; submitted 28.02.15; peer-reviewed by D Swendeman, S Michie; comments to author 02.04.15; revised version received 30.04.15; accepted 30.05.15; published 08.07.15.

Please cite as:
Mohr DC, Schueller SM, Riley WT, Brown CH, Cuijpers P, Duan N, Kwasny MJ, Stiles-Shields C, Cheung K
Trials of Intervention Principles: Evaluation Methods for Evolving Behavioral Intervention Technologies
J Med Internet Res 2015;17(7):e166
URL: http://www.jmir.org/2015/7/e166/
doi:10.2196/jmir.4391
PMID:26155878

©David C Mohr, Stephen M Schueller, William T Riley, C Hendricks Brown, Pim Cuijpers, Naihua Duan, Mary J Kwasny, Colleen Stiles-Shields, Ken Cheung. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 08.07.2015. 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

Mobile Phone Sensor Correlates of Depressive Symptom Severity in Daily-Life Behavior: An Exploratory Study

Sohrab Saeb[1,2], PhD; Mi Zhang[3], PhD; Christopher J Karr[1], MA; Stephen M Schueller[1], PhD; Marya E Corden[1], MPH; Konrad P Kording[2], PhD; David C Mohr[1], PhD

[1]Center for Behavioral Intervention Technologies, Department of Preventive Medicine, Northwestern University, Chicago, IL, United States
[2]Rehabilitation Institute of Chicago, Department of Physical Medicine and Rehabilitation, Northwestern University, Chicago, IL, United States
[3]Department of Electrical and Computer Engineering, Michigan State University, East Lansing, MI, United States

Corresponding Author:
David C Mohr, PhD
Center for Behavioral Intervention Technologies
Department of Preventive Medicine
Northwestern University
750 N Lake Shore Drive
Chicago, IL, 60611
United States
Phone: 1 312 503 1403
Fax: 1 312 908 9588
Email: d-mohr@northwestern.edu

Abstract

Background: Depression is a common, burdensome, often recurring mental health disorder that frequently goes undetected and untreated. Mobile phones are ubiquitous and have an increasingly large complement of sensors that can potentially be useful in monitoring behavioral patterns that might be indicative of depressive symptoms.

Objective: The objective of this study was to explore the detection of daily-life behavioral markers using mobile phone global positioning systems (GPS) and usage sensors, and their use in identifying depressive symptom severity.

Methods: A total of 40 adult participants were recruited from the general community to carry a mobile phone with a sensor data acquisition app (Purple Robot) for 2 weeks. Of these participants, 28 had sufficient sensor data received to conduct analysis. At the beginning of the 2-week period, participants completed a self-reported depression survey (PHQ-9). Behavioral features were developed and extracted from GPS location and phone usage data.

Results: A number of features from GPS data were related to depressive symptom severity, including circadian movement (regularity in 24-hour rhythm; r=-.63, P=.005), normalized entropy (mobility between favorite locations; r=-.58, P=.012), and location variance (GPS mobility independent of location; r=-.58, P=.012). Phone usage features, usage duration, and usage frequency were also correlated (r=.54, P=.011, and r=.52, P=.015, respectively). Using the normalized entropy feature and a classifier that distinguished participants with depressive symptoms (PHQ-9 score ≥5) from those without (PHQ-9 score <5), we achieved an accuracy of 86.5%. Furthermore, a regression model that used the same feature to estimate the participants' PHQ-9 scores obtained an average error of 23.5%.

Conclusions: Features extracted from mobile phone sensor data, including GPS and phone usage, provided behavioral markers that were strongly related to depressive symptom severity. While these findings must be replicated in a larger study among participants with confirmed clinical symptoms, they suggest that phone sensors offer numerous clinical opportunities, including continuous monitoring of at-risk populations with little patient burden and interventions that can provide just-in-time outreach.

(J Med Internet Res 2015;17(7):e175) doi:10.2196/jmir.4273

KEYWORDS

depression; mobile health (mHealth); activities of daily living; cluster analysis; classification

XSL•FO
RenderX

Introduction

Depression is a common mental disorder. Estimates of the 12-month prevalence rate for major depressive disorder range from 6.6-10.3%, while lifetime risk has been estimated at 16.6-17.1% [1-3]. Subthreshold depressive symptoms are far more common [4], causing significant impairment in people's lives and putting them at risk for future mental health concerns [5]. Depression, at both diagnosable and subthreshold levels, imposes a very high societal burden in terms of cost, lost productivity, morbidity, suffering, and mortality [6-8] and is a leading cause of disability and disease burden worldwide [9]. By the year 2020, the World Health Organization estimates that depression will be the second largest cause of "lost years of healthy life" worldwide [10].

Depression is treatable using a variety of methods, including antidepressants and psychotherapy; however, very few people who need treatment receive it [11]. It often takes months or years for depression to be identified and treated in our health care system, when it is treated at all [12]. One of the most common settings where depression is managed is primary care [13,14]; however, primary care physicians might fail to identify most patients with depressive symptoms [15,16]. Thus, more efficient methods of monitoring could significantly improve the delivery of services to those in need.

Mobile phones are becoming the most ubiquitous consumer device in our world. Equipped with powerful sensing, computation, and communication capabilities, mobile phones—specifically smartphones—can continuously monitor an individual's context including physical activity, location, and environment. Depression is associated with several behavioral components (eg, reduction in activity, psychomotor retardation, changes in sleep) and motivational states (eg, anhedonia), some of which may be detectable using mobile phone sensors [17,18]. Thus, mobile phones hold significant promise as a platform to monitor behavioral and environmental indicators of risk and resilience and to improve long-term management and treatment delivery to people suffering from depression.

Indeed, some work has shown promise in this area. A first study in our group found that phone sensor data could detect social patterns among depressed patients, but this was a small study with only 8 participants [19]. Other groups have found that phone sensors were effective at detecting social and sleep behaviors among patients with depression [20,21], and such features correlated significantly with severity of depressive symptoms [22].

The aim of this study was to extend previous work by focusing specifically on behavioral markers related to movement through geographic space, which we hypothesized would be related to depressive symptom severity, given depression results in decreased motivation, withdrawal, and activity. In addition, excessive use of mobile phones is considered compulsive behavior and has been linked to some symptoms of depression [23,24]. Thus, we also explored the relationship of depression symptom severity to the use of the phone that was used to collect the sensor data. To achieve these aims, we used our mobile phone app, Purple Robot [25,26], to collect global positioning system (GPS) location and phone usage data from participants with varying levels of depression severity. We defined a number of behavioral features based on these data and built classification and regression models to examine their relationship to depression symptom severity.

Methods

Participants

We recruited 40 adult participants from April-July 2013 using craigslist advertisements. Participants were eligible if they had an email account, computer, and broadband access to the Internet, were within a cellular network range the majority of the day, were able to speak and read English, were at least 19 years of age, and lived within the United States of America. Participants signed an online consent form, and research staff reviewed the consent over the phone. The study was approved by the Northwestern University Institutional Review Board.

At the beginning of the study, participants were asked to complete an online assessment consisting of a demographics questionnaire and the Patient Health Questionnaire-9 (PHQ-9), a commonly used measure of self-reported depressive symptom severity [27], which produces scores ranging from 0-27. Scores of less than 5 indicate no depression, 5-9 mild depression, 10-14 moderate depression, 15-19 moderately severe, and over 20 severe depression [28].

If participants owned and used an Android device with operating system 2.3 (Gingerbread) or higher, research staff assisted with the download, installation, and configuration of Purple Robot (see the Purple Robot section below). Participants who did not own a compatible phone were given an Android Nexus 4 with Purple Robot installed and configured. Phones were either picked up from the study's office or mailed directly to the participant.

Participants were instructed to keep the phone with them and charged throughout the day for 2 weeks. In addition, research staff explained that Purple Robot would be collecting GPS location and phone usage data. As part of the purpose of this study was to test Purple Robot's functionality in the field, research staff checked in periodically, via phone and email, with the participants to ensure the app was working properly and to answer questions.

Purple Robot

Purple Robot is an open-source Android app that we developed to collect mobile phone sensor data [25,26]. The app implements a store-and-forward architecture wherein the sensor data are gathered, stored on the device, and transmitted as network connectivity becomes available. This allows us to collect data in a variety of wireless connectivity scenarios with the confidence that intermittent network access did not affect the nature, quality, or quantity of the collected data.

Purple Robot anonymizes personally identifiable and other sensitive information before storage and transmission, using standard MD5 hashing and advanced encryption standard [29] algorithms. Once the data are anonymized, they are stored and

XSL•FO
RenderX

later transmitted to the data collection server before being deleted from the device. Sensor data residing on the server can be linked with other information gathered during the study only if the unique identifiers used by the participants and the study-specific keys used to encrypt the data are known.

The Purple Robot mobile app and supporting server infrastructure is capable of collecting information about the user's physical context (eg, motion), social settings (eg, number of Facebook friends), and phone usage behavior (eg, screen state). It also enables us to craft a complete data collection strategy configured for analyzing the relationship between depression and behavior data features of daily life.

In this study, we configured Purple Robot to collect the GPS location and phone usage data, as the aim of this study was to focus on behavioral markers related to movement through space and the phone usage behavior. In our next study, we plan to use Purple Robot to collect data from a variety of phone sensors.

Purple Robot sampled the GPS location sensor once every 5 minutes and collected phone usage data by detecting the screen on and off events.

Data Preprocessing

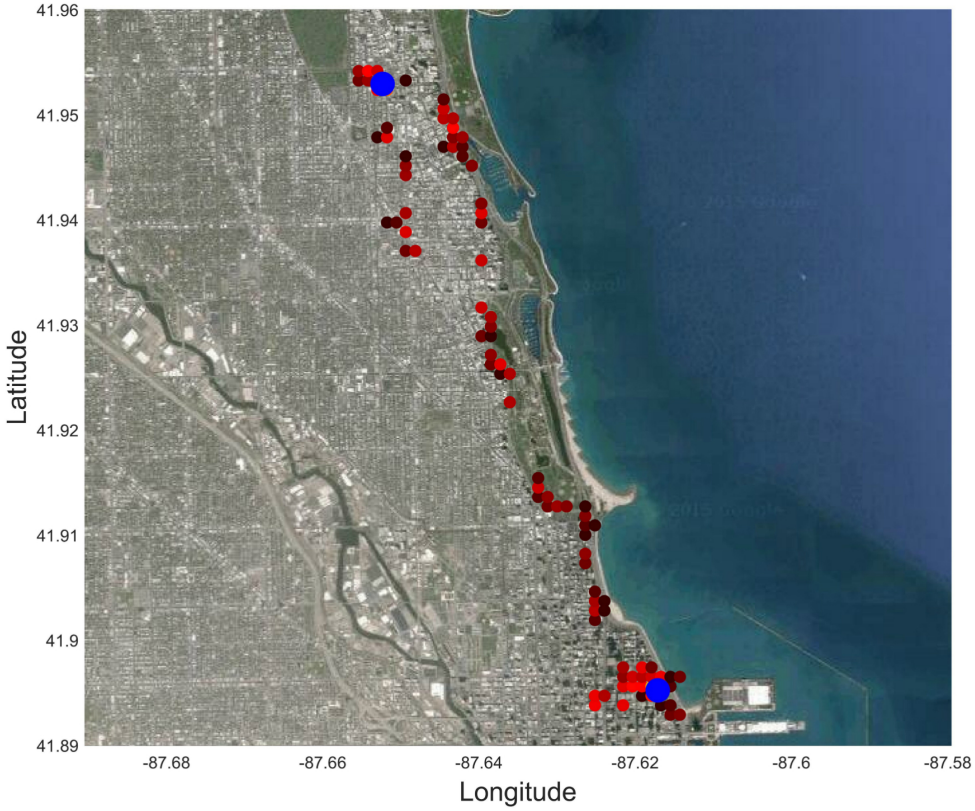
The goal of the data preprocessing stage was to facilitate the extraction of features from both the GPS location and the phone usage data.

For the location data (Figure 1), we used two procedures. The first procedure determined whether each GPS location data sample came from a stationary state (eg, working in an office) or a transition state (eg, walking on the street). To do so, we estimated the movement speed at each location data sample by calculating its time derivative and then used a threshold speed that defined the boundary between these two states. In this study, we set this threshold to 1 km/h.

The second procedure was clustering. We applied clustering only to the data samples in the stationary state. The goal was to identify the places where participants spent most of their time, such as home, workplaces, parks, etc. We used a distance-based clustering algorithm called *K*-means [30], in which the data were partitioned into *K* clusters such that the overall distances of the data points to the centers of their clusters were minimized. Because the number of clusters was unknown, we started with one cluster and increased the number of clusters until the distance of the farthest point in each cluster to its cluster center fell below a threshold. This threshold determined the maximum radius of a cluster, which was set to 500 meters in our study.

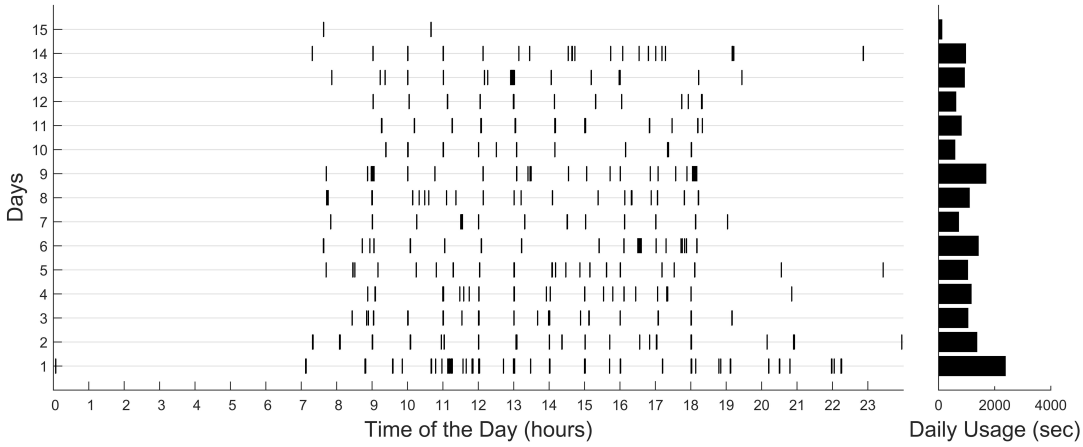
Phone usage data were gathered by looking at the periods of time when the phone screen was on (Figure 2). Given that the phone screen would go on when receiving notifications from apps such as text messages, we eliminated brief screen-on events not initiated by the participant that had durations of less than 30 seconds.

Figure 1. Example GPS location data, overlaid on satellite image. Each small circle represents a histogram bin, which has a size of 500 by 500 meters. The colors indicate the number of samples captured by each bin (brighter means more samples). The bigger blue circles show the center of the clusters detected by the clustering algorithm.



XSL•FO
RenderX

Figure 2. Example phone usage data from a participant. Each row is a day, and the black bars show the extent of time during which the phone has been is use. The bars on the right side show the overall phone usage duration for each day.



Feature Extraction

Location Variance

We defined location variance to measure the variability in a participant's GPS location. To calculate location variance, we used only the location data of stationary states (see Data Preprocessing). Specifically, location variance was computed as the logarithm of the sum of the statistical variances of the latitude and the longitude components of the location data:

$$\textit{Location Variance} = \log(\sigma_{lat}^2 + \sigma_{long}^2) \quad (1)$$

We applied the logarithm to compensate for the skewness in the distribution of location variance across participants.

Number of Clusters

Number of clusters represented the number of location clusters found by the *K*-means algorithm in the preprocessing stage.

Entropy

We defined entropy to measure the variability of the time the participant spent at the location clusters. This feature was developed based on the concept of entropy from information theory [31]. It was calculated as:

$$\textit{Entropy} = -\sum_i p_i \log p_i \quad (2)$$

where each i=1, 2, …, N represented a location cluster, N denoted the total number of location clusters, and p_i was the percentage of time the participant spent at the location cluster i. High entropy indicated that the participant spent time more uniformly across different location clusters, while lower entropy indicated greater inequality in the time spent across the clusters. For example, if a participant spent 80% of time at home and 20% at work, the resulting entropy would be −(0.8log0.8 + 0.2log0.2) ≈ 0.500, while if they spent 50% at home and 50% at work, the resulting entropy would be −(0.5log0.5 + 0.5log0.5) ≈ 0.693.

Normalized Entropy

We defined normalized entropy by dividing the entropy by its maximum value, which is the logarithm of the total number of clusters:

$$\textit{Normalized Entropy} = \textit{Entropy} / \log N \quad (3)$$

Unlike entropy, normalized entropy is invariant to the number of clusters and thus depends solely on the distribution of the visited location clusters. The value of normalized entropy ranges from 0-1, where 0 indicates that all location data points belong to the same cluster, and 1 implies that they are uniformly distributed across all the clusters.

Home Stay

Home stay measured the percentage of time a participant spent at home relative to other location clusters. To obtain this measure, we first needed to know which cluster represented the participant's home. We identified the home cluster based on two heuristics: (1) the home cluster is among the first to the third most visited clusters, and (2) the home cluster is the cluster most visited during the time period between 12 a.m. and 6 a.m. In our dataset, which did not contain participants having night shift work, these two heuristics led to one and only one location cluster for every participant.

Circadian Movement

We defined circadian movement to capture the temporal information of the location data. This feature measured to what extent the participants' sequence of locations followed a 24-hour, or circadian, rhythm. For example, if a participant left home for work and returned home from work around the same time each day, the circadian movement was high. On the contrary, a participant with a more irregular pattern of moving between locations had a lower circadian movement.

To calculate circadian movement, we first used the least-squares spectral analysis, also known as the Lomb-Scargle method [32], to obtain the spectrum of the GPS location data. Then, we calculated the amount of energy that fell into the frequency bins within a 24±0.5 hour period, in the following way:

$$E = \sum_i \text{psd}(f_i) / (i_1 - i_2) \quad (4)$$

where $i = i_1, i_1+1, i_1+2, \ldots, i_2$, and i_1 and i_2 represent the frequency bins corresponding to 24.5 and 23.5 hour periods. psd(f_i) denotes the power spectral density at each frequency bin f_i. We calculated E separately for longitude and latitude and obtained the total circadian movement as:

XSL•FO
RenderX

$CM = \log(E_{lat} + E_{long})$ (5)

We applied the logarithm to account for the skewness in the distribution.

Transition Time

Transition time represented the percentage of time during which a participant was in a non-stationary state (see Data Preprocessing). This was calculated by dividing the number of GPS location samples in transition states by the total number of samples.

Total Distance

Total distance measured the total distance in kilometers taken by a participant. It was calculated by accumulating the distances between the location samples.

Phone Usage Frequency

Phone usage frequency indicated, on average, how many times during a day a participant interacted with their phone.

Phone Usage Duration

Phone usage duration measured, on average, the total time in seconds that a participant spent each day interacting with their phone.

Relationship Between Features and Levels of Depression

We performed a preliminary statistical analysis to find out how each feature would correspond to levels of depressive symptoms and whether it would be able to distinguish people with any level of depression from those with none. The former was investigated by correlating each feature with the PHQ-9 score that was obtained at the beginning of the study. The latter was explored first by dividing participants into those with depressive symptoms (PHQ-9 ≥5) and the ones without (PHQ-9 <5). The cutoff of 5 on the PHQ-9 score was used because scores in this range are indicative of "no symptoms" of depression and those who reach this range after treatment are considered to be in full remission.

Estimating Depression States from Features

Score Estimation Model

We used a linear regression model to estimate each participant's PHQ-9 score using the features extracted from their phone sensor data. The model was defined as the following:

$Depression\ Score = a_0 + a_i F_i + a_2 F_2 + \ldots + a_n F_n$ (6)

where n is the number of features. The coefficients $a_0, a_1, \ldots a_n$ were obtained by minimizing the squared error between the estimated and the true PHQ-9 scores (see Model Optimization).

Classification Model

We used a logistic regression classifier to identify participants who had symptoms of depression (PHQ-9 ≥5) versus the ones with no symptoms (PHQ-9 <5). This classifier consisted of a linear model and a logistic sigmoid function, $g(x) = 1 / (1+\exp(-x))$, that generated values between 0 and 1 indicating the probability of the participant having depressive symptoms:

$P(Depressive\ Symptoms) = g(b_0 + b_1F_1 + b_2F_2 + \ldots + b_nF_n)$ (7)

When P(*Depressive Symptoms*) was higher than 0.5, we considered the participant to have depressive symptoms. Otherwise, we considered there was no sign of depression.

We used an optimization procedure to adjust the parameters b_0, $b_1, \ldots, b_n$ (see Model Optimization).

Model Optimization

We used the least squares approach to adjust the parameters of both score estimation and classification models. This method performs well as long as the number of features relative to the number of samples is low. Otherwise, the model overfits the data. To minimize the overfitting problem, we used the elastic-net regularization method when the number of features was high.

The elastic-net regularization prevents the coefficient from becoming too large by adding the following penalizing term to the cost function:

$H(K,\lambda_1,\lambda_2) = \lambda_1 ||K||_1 + \lambda_1 ||K||_2$ (8)

where $K=k_0, k_1, \ldots, k_n$ is the vector containing the regression or the classification model parameters, and $\| K \|_1 = \sum_i |k_i|$ and $\| K \|_2 = \sqrt{\sum_i k_i^2}$ are its first (L1) and second (L2) norms (i=1,2,…,n). The coefficients λ_1 and λ_2 are optimized by cross-validating on the training data. Elastic-net regularization has been shown to outperform other regularization methods especially in cases where some of the features are strongly correlated [33].

Model Evaluation

To evaluate regression and classification models, we created 1000 bootstrapped sets of features and their corresponding PHQ-9 scores. Then we trained and cross-validated the models on each set using the leave-one-participant-out method.

To assess the performance of each score estimation model, we calculated the normalized root mean square deviation (NRMSD), which measures the percent difference between the PHQ-9 scores estimated by the model on the test participants and their true scores. We used the observed range of PHQ-9 scores, which was 0-17, to normalize the NRMSDs. To evaluate the performance of each classification model, we evaluated its accuracy, sensitivity, and specificity in identifying participants with depressive symptoms as compared to the ground truth.

Results

Participant Characteristics and Adherence

All 40 participants completed the 2 weeks of the study. However, due to insufficient sensor data for a number of participants, we considered 28 of them for the data analysis. Among these, 18 were considered for GPS location and 21 for phone usage data analysis. These two analyses were performed independently.

The 12 excluded participants did not provide sufficient GPS location and/or phone usage data for our analysis, meaning that

XSL•FO
RenderX

their data were not available for more than 50% of the time. The reasons for unavailability of data included problems in the connection between the early version of Purple Robot and the data server, patients not charging their phones, and unavailability of any network connections for a long period of time for certain locations.

The 28 participants included in the analysis were 20 females and 8 males. Their ages ranged from 19-58, with a mean of 28.9 (SD 10.1). Their PHQ-9 scores ranged from 0-17, with a mean of 5.57 (SD 4.9). Of the 28 participants, 14 (50%) had no signs of depression (PHQ-9 <5) with an average PHQ-9 score of 1.5 (SD 1.34), and the other 14 participants (50%) were in the mild to severe range (PHQ-9 ≥5), with an average PHQ-9 score of 9.64 (SD=3.54). Participants who were not included in any analyses due to inadequate data were not statistically different from these 28 participants in their age, gender, or PHQ-9 scores.

Among the participants who were considered for location data analysis, 9 had depressive symptoms and 9 did not. For the ones considered for phone usage data analysis, these numbers were 10 and 11, respectively.

Relationship Between Features and Levels of Depression

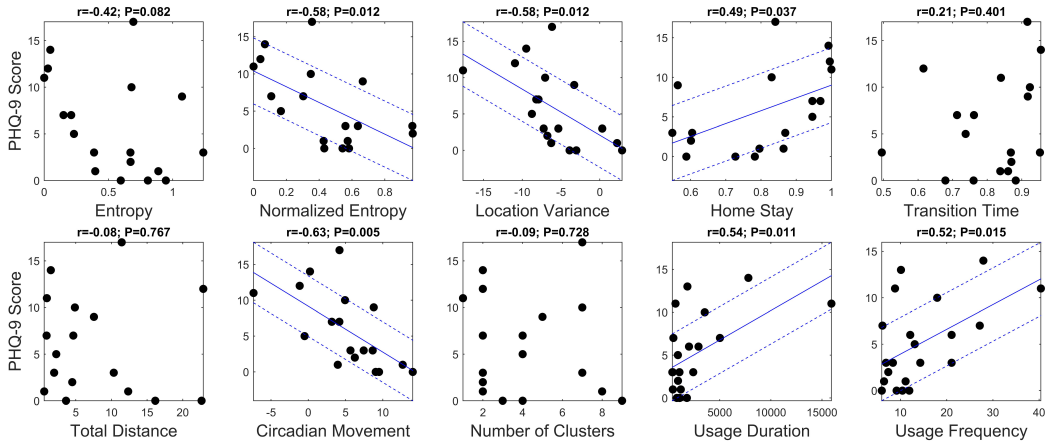
GPS location and phone usage sensor features were calculated as described in Feature Extraction. The number of location clusters that was found by the *K*-means algorithm ranged from 1-9, with an average of 4.06. The average daily phone usage duration across the participants was about 41 minutes (SD 57 minutes) with an average daily usage frequency of 14.2 times (SD 8.69).

The correlation analysis between the features and the PHQ-9 scores revealed that 6 of the 10 features were significantly correlated to the scores (Figure 3). Specifically, circadian movement, normalized entropy, and location variance showed strong correlations with Pearson's correlation coefficients of -.63, -.58, and -.58, respectively. Both phone usage features, usage duration and usage frequency, were also significantly correlated with r=-.54 and 0.52, respectively.

The *t* tests between participants with depressive symptoms and the ones without (Figure 4) also revealed that the value of the same six features (circadian movement, normalized entropy, location variance, home stay, phone usage duration, and phone usage frequency) were significantly different between the participants with no sign of depression (PHQ-9 <5) and the rest (PHQ-9 ≥5).

A correlation analysis across the features revealed that a number of them were highly correlated (Figure 5). Noticeably, there was a significant correlation between normalized entropy, location variance, and home stay. This is not surprising, as all these variables measure the amount of movement through space in different ways. However, the significant correlation between circadian movement and location variance is interesting and indicates that participants with more mobility also had more regular patterns of movement. The correlation between phone usage duration and frequency was also high (r=.89).

Figure 3. Scatter plots for location and phone usage data versus PHQ-9 scores, respectively. The coefficient of correlation between each feature and PHQ-9 scores and its corresponding P-value is shown on top of each plot. Solid and dashed lines, shown only for strong correlations (P<.05), show the fitted regression model and +/- root mean square deviation from the model, respectively.



XSL•FO
RenderX

Figure 4. Comparison of location and usage feature statistics between participants with no symptoms of depression (blue) and the ones with (red). Feature values are scaled between 0 and 1 for easier comparison. Boxes extend between 25th and 75th percentiles, and whiskers show the range. Horizontal solid lines inside the boxes are medians. One, two, and three asterisks show significant differences at $P<.05$, $P<.01$, and $P<.001$ levels, respectively (ENT, entropy; ENTN, normalized entropy; LV, location variance; HS, home stay; TT, transition time; TD, total distance; CM, circadian movement; NC, number of clusters; UF, usage frequency; UD, usage duration).

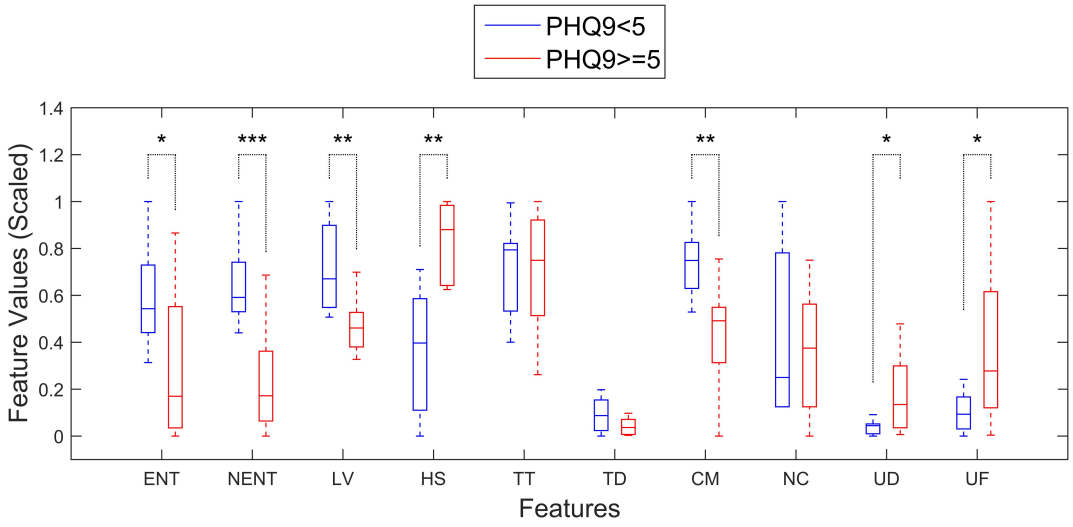
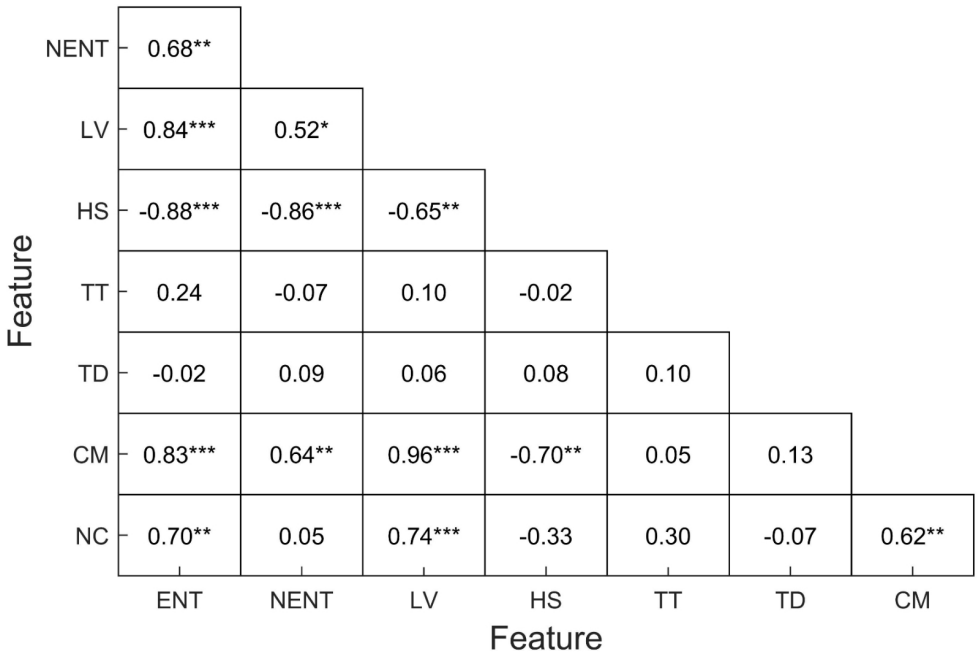


Figure 5. Coefficients of correlation between location features. One, two, and three asterisks indicate significant correlation levels at P<.05, P<.01, and P<.001, respectively (ENT, entropy; ENTN, normalized entropy; LV, location variance; HS, home stay; TT, transition time; TD, total distance; CM, circadian movement; NC, number of clusters).



Estimating Depression States from Location Features

The results of the earlier statistical analysis suggested that we may be able to estimate an unseen subject's depression state using some of our features. To test this hypothesis, we trained and cross-validated score prediction and classification models (Equations 6-7) using the procedure described in Model Evaluation.

We trained the models first on each individual feature and then all features combined together. As the results (Table 1) show, the models trained on the features that had stronger correlations with PHQ-9 scores were better able to distinguish the participants with depressive symptoms from those who had none. Columns 2-4 show the cross-validated accuracy, sensitivity, and specificity of each classification model (Equation 7) in classifying participants into the ones with and without depressive symptoms. Column 5 shows the cross-validated NRMSDs of the PHQ-9 score estimation models (Equation 6). Specifically, the four features normalized entropy, location variance, home stay, and circadian movement achieved the lowest NRMSDs and highest accuracies. These performances, however, did not improve by combining the features. This can be the result of some unavoidable overfitting as the number of input variables increases, which leads to a poor generalization.

XSL•FO
RenderX

Table 1. Classification of participants with and without depressive symptoms and estimating their PHQ-9 scores using location features individually and aggregated.

	Classification (PHQ9<5 vs PHQ9≥5)			PHQ-9 score estimation
Training features	% mean accuracy (SD)	% mean sensitivity	% mean specificity	Mean NRMSD (SD)
Entropy	69.7 (3.5)	66.8	72.7	0.262 (0.017)
Normalized entropy	86.5 (3.4)	88.4	84.9	0.235 (0.016)
Location variance	75.7 (4.6)	80.2	71.5	0.229 (0.014)
Home stay	75.9 (4.9)	80.5	71.7	0.253 (0.015)
Transition time	41.1 (9.2)	43.4	38.7	0.303 (0.020)
Total distance	56.4 (6.6)	69.6	43.4	0.343 (0.041)
Circadian movement	78.6 (4.1)	80.1	77.5	0.222 (0.014)
Number of clusters	41.5 (8.9)	47.4	35.5	0.305 (0.022)
All	78.8 (6.2)	83.6	74.5	0.251 (0.023)

Estimating Depression States from Phone Usage Features

We performed the same analyses on the phone usage features. Since the number of these features (n=2) was much smaller than the number of samples (n=21), both score prediction and classification models could be directly applied to the combined feature space without using elastic-net regularization. The results (Table 2) show that each of the usage frequency and usage duration features could provide acceptable accuracies and NRMSDs without further improvement by aggregating them. Columns 2-4 show the cross-validated accuracy, sensitivity, and specificity of each classification model (Equation 7) in classifying participants into the ones with and without depressive symptoms. Column 5 shows the cross-validated NRMSDs of the PHQ-9 score estimation models (Equation 6).

Table 2. Classification of participants with and without depressive symptoms and estimating their PHQ-9 scores using phone usage features individually and aggregated.

	Classification (PHQ9<5 vs PHQ9≥5)			PHQ9 score estimation
Training features	% mean accuracy (SD)	% mean sensitivity	% mean specificity	Mean NRMSD (SD)
Usage duration	74.2 (3.4)	64.0	83.9	0.268 (0.018)
Usage frequency	68.6 (4.1)	56.4	79.6	0.249 (0.013)
All	65.7 (4.9)	55.7	74.9	0.273 (0.019)

Discussion

Principal Findings

This study reported on the potential to use commonly available mobile phone sensor data, including GPS and phone usage, to identify depressive symptom severity. We extracted a number of semantically meaningful features from these data and found a strong correlation between a number of them and the PHQ-9 scores. These features included normalized entropy, location variance, home stay, circadian movement, and phone usage duration and frequency. By training score estimation models on each of these six features, we could estimate the PHQ-9 scores of unseen participants with a relatively low error (NRMSD). In addition, classifiers trained on these features were able to discriminate between those with and those without symptoms with a high degree of accuracy, good sensitivity, and specificity.

The normalized entropy feature measured the frequency with which a person visited different locations and the distribution of that frequency across locations. The high negative correlation that was found between this feature and the PHQ-9 scores indicated that people with greater depressive symptom severity visited fewer locations and were more likely to favor some locations over others. Part of this was likely due to the increased amount of time people with depressive symptoms spent at home, measured by the home stay feature. The finding for the location variance feature, on the other hand, indicated that people with depressive symptoms tend to move less through geographic space. These findings are consistent with the patterns of loss of motivation, decreased activity, and social withdrawal that characterize depression [17,18].

The finding for the circadian rhythm of movement through space fits into well-established literature investigating alterations to the wake-sleep patterns associated with depression [34,35]. These data suggest that disruptions in behavioral patterns during waking hours include not only the volume of activity but may also extend to the patterns of behavior. These pattern changes may be due to the genetic and hormonal factors [36] that have been implicated in depression-related circadian rhythm changes or may be a result of low motivation and decreased organization.

Phone usage data were also strongly correlated to depressive symptom severity. Greater levels of depressive symptom severity were related to greater phone usage duration and frequency. This observation is supported by a number of

XSL•FO
RenderX

previous studies that have found a correlation between mobile phone use and some depressive symptoms [23,24]. However, we should note that phone usage in this context was defined as any interaction with the phone, and we were not able to isolate the specific types of interactions (eg, using apps, texting). Thus, it is difficult to determine which specific behaviors were related to symptom severity.

Limitations

While we believe that our study has revealed some of the daily-life correlates of depression that can be captured by mobile phones, the results are very preliminary, and a number of caveats must be mentioned. First, this study examined only the association between self-reported depressive symptoms and features derived from location and phone usage data. Thus, we cannot infer any causal relationship here. In fact, while the PHQ-9 is a well-validated measure of depression, we cannot exclude the possibility that factors other than depressive symptoms are responsible for these relationships. For example, the results may be due to other unmeasured factors, such as chronic illness or dispositional factors, which result both in depressive symptoms and differences in behavioral patterns.

Second, while some participants demonstrated levels of depressive symptoms consistent with clinical levels of depression, this was a small sample that was not necessarily representative of typical trends seen in people with depression. Future research could recruit more representative participants with depression and match them on characteristics that might impact one's pattern of movements through geographic locations (eg, occupational status, size of social network, or chronic health problems).

Finally, we did not attempt to correct for the possible effect of multiple comparison. However, given our interest in exploring potential indicators of depressive symptoms, the increased likelihood of Type II errors introduced by such corrections might undermine important features. A major goal of such corrections is to increase confidence in one's findings and given the preliminary nature of our results, we urge future efforts to cross-validate these relationships in larger-scale investigations. Nevertheless, we believe the computation of behaviorally meaningful features (eg, normalized entropy, circadian movement) and the relationship of these features with depression found in this study might provide a valuable starting place for subsequent investigations of the use of sensor data for the monitoring and the detection of depression.

Conclusions

Regardless of these limitations, the ability to passively detect behavioral factors related to depression, such as activity levels and their patterns, opens the possibility of a new generation of behavioral intervention technologies that can passively detect and positively reinforce behaviors that are likely to improve depression (eg, engagement in activities that provide pleasure, a sense of accomplishment, or involve social engagement) and offer support when risk states are detected (eg, withdrawal, staying at home). This can improve the identification of depression and the ability of health care settings to allocate resources to those in need and overcome the individual and systemic barriers to conventional psychological treatment [37].

The use of phone sensors allows the capture of information that is potentially indicative of depressive symptoms without the use of questionnaires or requiring the person to use special devices. Phones fit into the fabric of people's lives. People tend to keep phones with them all or most of the time, and phones can provide data unobtrusively and with no effort on the part of the user. This capacity offers new opportunities to identify human behavior patterns associated with depression or other health and mental health disorders. Furthermore, behavioral features might be more sensitive to changes in a person's daily life that indicate early benefit from treatment or highlight potential areas for improvement. As such, as these features and the link between them and depression become better understood, they may play an important role in understanding the progression of depression and its response to treatment.

Acknowledgments

This research was funded by research grant P20 MH090318 from the US National Institute of Mental Health. Author SMS was supported by a grant from the National Institute of Mental Health K08 MH 102336.

Conflicts of Interest

None declared.

References

1. Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, National Comorbidity Survey Replication. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). JAMA 2003 Jun 18;289(23):3095-3105. [doi: 10.1001/jama.289.23.3095] [Medline: 12813115]
2. Kessler RC, McGonagle KA, Zhao S, Nelson CB, Hughes M, Eshleman S, et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. Arch Gen Psychiatry 1994 Jan;51(1):8-19. [Medline: 8279933]
3. 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]

XSL•FO
RenderX

4. Karsten J, Hartman CA, Smit JH, Zitman FG, Beekman AT, Cuijpers P, et al. Psychiatric history and subthreshold symptoms as predictors of the occurrence of depressive or anxiety disorder within 2 years. Br J Psychiatry 2011 Mar;198(3):206-212 [FREE Full text] [doi: 10.1192/bjp.bp.110.080572] [Medline: 21357879]
5. Lewinsohn PM, Solomon A, Seeley JR, Zeiss A. Clinical implications of "subthreshold" depressive symptoms. J Abnorm Psychol 2000 May;109(2):345-351. [Medline: 10895574]
6. Wells KB, Miranda J, Bauer MS, Bruce ML, Durham M, Escobar J, et al. Overcoming barriers to reducing the burden of affective disorders. Biol Psychiatry 2002 Sep 15;52(6):655-675. [Medline: 12361673]
7. Wang PS, Simon G, Kessler RC. The economic burden of depression and the cost-effectiveness of treatment. Int J Methods Psychiatr Res 2003;12(1):22-33. [Medline: 12830307]
8. Martin JK, Blum TC, Beach SR, Roman PM. Subclinical depression and performance at work. Soc Psychiatry Psychiatr Epidemiol 1996 Jan;31(1):3-9. [Medline: 8821918]
9. Ferrari AJ, Charlson FJ, Norman RE, Patten SB, Freedman G, Murray CJL, et al. Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010. PLoS Med 2013 Nov;10(11):e1001547 [FREE Full text] [doi: 10.1371/journal.pmed.1001547] [Medline: 24223526]
10. World Health Organization (Geneva). Mental Health: A Call to Action by World Health Ministers. 2001. URL: http://www.who.int/mental_health/media/en/249.pdf [accessed 2015-07-09] [WebCite Cache ID 6ZtjMXm7j]
11. Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005 Jun;62(6):617-627 [FREE Full text] [doi: 10.1001/archpsyc.62.6.617] [Medline: 15939839]
12. Wang PS, Berglund P, Olfson M, Pincus HA, Wells KB, Kessler RC. Failure and delay in initial treatment contact after first onset of mental disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005 Jun;62(6):603-613. [doi: 10.1001/archpsyc.62.6.603] [Medline: 15939838]
13. Simon GE, VonKorff M. Recognition, management, and outcomes of depression in primary care. Arch Fam Med 1995 Feb;4(2):99-105. [Medline: 7842160]
14. van Weel C, van Weel-Baumgarten E, van Rijswijk E. Treatment of depression in primary care. BMJ 2009;338:b934. [Medline: 19299476]
15. Schulberg HC, Saul M, McClelland M, Ganguli M, Christy W, Frank R. Assessing depression in primary medical and psychiatric practices. Arch Gen Psychiatry 1985 Dec;42(12):1164-1170. [Medline: 4074109]
16. Wells KB, Hays RD, Burnam MA, Rogers W, Greenfield S, Ware JE. Detection of depressive disorder for patients receiving prepaid or fee-for-service care. Results from the Medical Outcomes Study. JAMA 1989 Dec 15;262(23):3298-3302. [Medline: 2585674]
17. Prigerson HG, Monk TH, Reynolds CF, Begley A, Houck PR, Bierhals AJ, et al. Lifestyle regularity and activity level as protective factors against bereavement-related depression in late-life. Depression 1995;3(6):297-302. [doi: 10.1002/depr.3050030607]
18. Vallée J, Cadot E, Roustit C, Parizot I, Chauvin P. The role of daily mobility in mental health inequalities: the interactive influence of activity space and neighbourhood of residence on depression. Soc Sci Med 2011 Oct;73(8):1133-1144. [doi: 10.1016/j.socscimed.2011.08.009] [Medline: 21903318]
19. Burns MN, Begale M, Duffecy J, Gergle D, Karr CJ, Giangrande E, et al. Harnessing context sensing to develop a mobile intervention for depression. J Med Internet Res 2011;13(3):e55 [FREE Full text] [doi: 10.2196/jmir.1838] [Medline: 21840837]
20. Doryab A, Min J, Wiese J, Zimmerman J, Hong J. Detection of behavior change in people with depression. : AAAI; 2014 Presented at: AAAI Workshop on Modern Artificial Intelligence for Health Analytics; 2014; Québec City, Québec, Canada.
21. Rabbi M, Ali S, Choudhury T, Berke E. Passive and in-situ assessment of mental and physical well-being using mobile sensors. : ACM; 2011 Presented at: International Conference on Ubiquitous Computing (UbiComp); 2011; Beijing, China p. 385-394.
22. Wang R, Chen F, Chen Z, Li T, Harari G, Tignor S. StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones. : ACM; 2014 Presented at: International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp); 2014; Seattle, Washington, USA p. 3-14.
23. Thomée S, Härenstam A, Hagberg M. Mobile phone use and stress, sleep disturbances, and symptoms of depression among young adults - a prospective cohort study. BMC Public Health 2011;11:66.
24. Lee YK, Chang CT, Lin Y, Cheng ZH. The dark side of smartphone usage: Psychological traits, compulsive behavior and technostress. Computers in Human Behavior 2014 Feb;31:373-383. [doi: 10.1016/j.chb.2013.10.047]
25. CBITs TECH Website, PurpleRobot. URL: http://tech.cbits.northwestern.edu/purple-robot/ [accessed 2015-01-20] [WebCite Cache ID 6Vja2Kh1a]
26. Schueller SM, Begale M, Penedo FJ, Mohr DC. Purple: a modular system for developing and deploying behavioral intervention technologies. J Med Internet Res 2014;16(7):e181 [FREE Full text] [doi: 10.2196/jmir.3376] [Medline: 25079298]
27. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [Medline: 11556941]

XSL•FO
RenderX

28. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-5). Washington, DC: American Psychiatric Association; 2013.
29. Heron S. Advanced Encryption Standard (AES). Network Security 2009 Dec;2009(12):8-12. [doi: 10.1016/S1353-4858(10)70006-4]
30. Arthur D, Vassilvitskii S. K-means++: the advantages of careful seeding. 2007 Presented at: Annual ACM-SIAM Symposium on Discrete Algorithms; 2007; Philadelphia, PA, USA p. 1027-1035.
31. Shenkin PS, Erman B, Mastrandrea LD. Information-theoretical entropy as a measure of sequence variability. Proteins 1991;11(4):297-313. [doi: 10.1002/prot.340110408] [Medline: 1758884]
32. Press WH. Numerical Recipes 3rd Edition: The Art of Scientific Computing. NY, USA: Cambridge University Press; 2007.
33. Zou H, Hastie T. Regularization and variable selection via the elastic net. J Royal Statistical Soc B 2005 Apr;67(2):301-320. [doi: 10.1111/j.1467-9868.2005.00503.x]
34. Kronfeld-Schor N, Einat H. Circadian rhythms and depression: human psychopathology and animal models. Neuropharmacology 2012 Jan;62(1):101-114. [doi: 10.1016/j.neuropharm.2011.08.020] [Medline: 21871466]
35. Raoux N, Benoit O, Dantchev N, Denise P, Franc B, Allilaire JF, et al. Circadian pattern of motor activity in major depressed patients undergoing antidepressant therapy: relationship between actigraphic measures and clinical course. Psychiatry Res 1994 Apr;52(1):85-98. [Medline: 8047624]
36. Wirz-Justice A. Biological rhythm disturbances in mood disorders. Int Clin Psychopharmacol 2006 Feb;21 Suppl 1:S11-S15. [doi: 10.1097/01.yic.0000195660.37267.cf] [Medline: 16436934]
37. Mohr DC, Ho J, Duffecy J, Baron KG, Lehman KA, Jin L, et al. Perceived barriers to psychological treatments and their relationship to depression. J Clin Psychol 2010 Apr;66(4):394-409 [FREE Full text] [doi: 10.1002/jclp.20659] [Medline: 20127795]

Abbreviations

GPS: global positioning system
NRMSD: normalized root mean square deviation
PHQ-9: Patient Health Questionnaire-9

Edited by G Eysenbach; submitted 21.01.15; peer-reviewed by J Ruwaard, J Bardram; comments to author 26.03.15; revised version received 12.05.15; accepted 21.06.15; published 15.07.15.

Please cite as:
Saeb S, Zhang M, Karr CJ, Schueller SM, Corden ME, Kording KP, Mohr DC
Mobile Phone Sensor Correlates of Depressive Symptom Severity in Daily-Life Behavior: An Exploratory Study
J Med Internet Res 2015;17(7):e175
URL: http://www.jmir.org/2015/7/e175/
doi:10.2196/jmir.4273
PMID:26180009

©Sohrab Saeb, Mi Zhang, Christopher J Karr, Stephen M Schueller, Marya E Corden, Konrad P Kording, David C Mohr. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 15.07.2015. 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 Web-Based Osteoarthritis Management Resource My Joint Pain Improves Quality of Care: A Quasi-Experimental Study

Hema Umapathy[1], BBiomedSc, MBiomedSc; Kim Bennell[2], PhD; Chris Dickson[3], BA (Hons); Fiona Dobson[2], PhD; Marlene Fransen[4], PhD; Graeme Jones[5], PhD; David J Hunter[1,6], MBBS, PhD, FRACP

[1]Institute of Bone and Joint Research, The Kolling Institute, Department of Rheumatology, University of Sydney, Sydney, Australia

[2]Centre for Health, Exercise and Sports Medicine, Department of Physiotherapy, University of Melbourne, Melbourne, Australia

[3]Arthritis Australia, Sydney, Australia

[4]Discipline of Physiotherapy, Clinical and Rehabilitation Sciences Research Group, Faculty of Health Sciences, University of Sydney, Lidcombe, Australia

[5]Menzies Research Institute Tasmania, University of Tasmania, Hobart, Australia

[6]Rheumatology Department, Royal North Shore Hospital, Sydney, Australia

Corresponding Author:
David J Hunter, MBBS, PhD, FRACP
Institute of Bone and Joint Research, The Kolling Institute
Department of Rheumatology
University of Sydney
Clinical administration 7A, Royal North Shore Hospital
Reserve Road
Sydney, 2065
Australia
Phone: 61 29463 1896
Fax: 61 29463 1077
Email: david.hunter@sydney.edu.au

Abstract

Background: Despite the availability of evidence-based guidelines for conservative treatment of osteoarthritis (OA), management is often confined to the use of analgesics and waiting for eventual total joint replacement. This suggests a gap in knowledge for persons with OA regarding the many different treatments available to them.

Objective: Our objective was to evaluate outcomes after usage of a Web-based resource called My Joint Pain that contains tailored, evidence-based information and tools aimed to improve self-management of OA on self-management and change in knowledge.

Methods: A quasi-experimental design was used to evaluate the My Joint Pain website intervention over a 12-month period. The intervention provided participants with general and user-specific information, monthly assessments with validated instruments, and progress-tracking tools. A nationwide convenience sample of 195 participants with self-assessed hip and/or knee OA completed both baseline and 12-month questionnaires (users: n=104; nonusers: n=91). The primary outcome measure was the Health Evaluation Impact Questionnaire (heiQ) to evaluate 8 different domains (health-directed activity, positive and active engagement in life, emotional distress, self-monitoring and insight, constructive attitudes and approaches, skill and technique acquisition, social integration and support, health service navigation) and the secondary outcome measure was the 17-item Osteoarthritis Quality Indicator (OAQI) questionnaire to evaluate the change in appropriateness of care received by participants. Independent *t* tests were used to compare changes between groups for the heiQ and chi-square tests to identify changes within and between groups from baseline to 12 months for each OAQI item.

Results: Baseline demographics between groups were similar for gender (152/195, 77.9% female), age (mean 60, SD 9 years) and body mass index (mean 31.1, SD 6.8 kg/m^2). With the exception of health service navigation, mean effect sizes from all other heiQ domains showed a positive trend for My Joint Pain users compared to the nonusers, although the differences between groups did not reach statistical significance. Within-group changes also showed improvements among the users of the My Joint Pain website for self-management (absolute change score=15%, *P*=.03), lifestyle (absolute change score=16%, *P*=.02), and physical activity (absolute change score=11%, *P*=.04), with no significant improvements for the nonusers. Following 12 months of exposure

XSL•FO
RenderX

to the website, there were significant improvements for users compared to nonusers in self-management (absolute change score 15% vs 2%, *P*=.001) and weight reduction (absolute change scores 3% vs –6%, *P*=.03) measured on the OAQI.

Conclusions: The My Joint Pain Web resource does not significantly improve overall heiQ, but does improve other important aspects of quality of care in people with hip and/or knee OA. Further work is required to improve engagement with the website and the quality of information delivered in order to provide a greater impact.

(J Med Internet Res 2015;17(7):e167) doi:10.2196/jmir.4376

KEYWORDS

quality of health care; self-care; osteoarthritis; eHealth; Internet

Introduction

Osteoarthritis (OA) is a highly prevalent and disabling disease and the leading cause of chronic pain [1,2]. It is estimated to affect 1 in 8 adults and the risk for mobility disability due to knee OA is unsurpassed by any other condition in those older than age 65 years [1,3,4]. The aging population and increasing rates of obesity contribute to the estimate that by 2020, the prevalence of OA will have doubled [5].

Numerous evidence-based recommendations for the care of OA that advocate conservative nonpharmacological treatments as the cornerstone of management have been produced to guide health professionals in OA management [6-11]. However, current clinical practice does not reflect these core recommendations and treatment is often limited to analgesics and eventual joint replacement [12-14]. Furthermore, an underuse of key conservative treatments and low levels of referrals to allied health professionals who may provide key lifestyle interventions was reported both internationally and within Australia [12,15].

Surveys of patients with OA reveal a large degree of dissatisfaction with their treatment and a preference for conservative treatment rather than surgery; 81% of patients have indicated that they would not accept surgery if offered and patients with a preference for a treatment most commonly choose physiotherapy [16,17]. The lack of efficacy of current practice coupled with patient dissatisfaction with current care attests to a large and urgent need for people with OA to receive appropriate evidence-based information about treatment options. This need is further supported by patient dissatisfaction with information received from health care professionals, low levels of knowledge regarding OA, and the low levels of patient involvement in treatment decisions [18-20].

The provision of evidence-based information to patients outside the clinical encounter has been long carried through patient education and self-management programs with an aim to result in self-efficacy. Self-management allows patients to recognize problems and play an active role toward improving their condition, essentially being their own caregiver [21].

Consumer-focused strategies are essential for chronic disease self-management to improve outcomes. Although education is a core component of treatment, a 2014 Cochrane review found that current self-management programs for OA offer little benefit and recommended investigation of alternate models of delivery [22]. The potential utility of eHealth for delivery of consumer-focused strategies is increasingly being recognized in-line with the rapid explosion of such technology across all socioeconomic levels and ages [23,24]. It has been suggested in a consensus rating that eHealth interventions for chronic disease management may have greater reach, adoption, implementation, and long-term effectiveness over traditional face-to-face intervention modalities and thereby greater public health impact [25].

This study reports several outcomes following the use of a publicly available Web-based resource, My Joint Pain, over a 12-month period among participants with hip and/or knee OA. My Joint Pain was designed to be a freely available hub of evidence-based information and self-management support resources. The website aims to empower and improve informed treatment decision making in the health care setting. As such, the aim of this study was to evaluate outcomes in users of My Joint Pain on the quality of care and self-management in people with hip and/or knee OA.

Methods

Study Design

The evaluation of My Joint Pain [26] was a quasi-experimental study that involved a group of users of the website and a group of nonusers with OA of the hip and/or knee. Outcome measures were administered at baseline and at 12-month follow-up with pre-post changes within each group and a comparison of changes between groups evaluated. The study was approved by the Human Research Ethics Committee of the Universities of Melbourne and Sydney and all participants provided online informed consent.

Participants

A nationwide convenience sample of adults aged 50 years and older with self-assessed OA in at least one hip or knee joint, access to the Internet, and an email account were recruited using an online strategy involving advertisements placed on various websites, including Arthritis Australia; Melbourne Physiotherapy Department; Centre for Health, Exercise and Sports Medicine; and the Sydney Medical School. Interested participants were directed to a validated screening tool to assess their eligibility. The screening tool contained a series of weighted questions that produced a likelihood of OA score for a nominated hip or knee joint [27]. Participants who received a score above a minimum risk of OA were eligible to participate.

Procedure

All communication with participants occurred via email and online. Eligible participants completed baseline surveys to

XSL•FO
RenderX

collect outcome measures as well as basic demographics and health services utilization. Participants were then informed the My Joint Pain website was accessible for use if they chose to use it. Participants were then contacted 12 months later to complete follow-up surveys. Additionally at follow-up, participants were required to indicate if they were aware of and had used the My Joint Pain website in the previous 12 months. Participants who indicated usage of the website were classified as the intervention group, whereas all other respondents were classified as nonusers.

Intervention

An expert content committee composed of leading OA researchers, health professionals, and consumers guided Arthritis Australia in the development of My Joint Pain. The website's functionality was informed by research outlining the key elements required for effective management of chronic disease. Criteria for judging the quality of patient decision aids, prepared by the International Patient Decision Aids Standards (IPDAS) Collaboration also helped to provide a functional framework for the site [28]. Health information was reviewed in-line with Australia's National Health & Medical Research Council (NHMRC) Guidelines for Consumer Information. The website also complied with the Health on the Net (HON) code standard, an ethical certification for quality health information on the Internet, developed by the HON Foundation, a nongovernmental organization.

The aim of the My Joint Pain website was to provide a credible, tailored information source with a variety of self-assessment tools to improve disease knowledge and self-management of OA. The resource was developed in collaboration with Arthritis Australia (a charitable not-for-profit organization and Australia's peak arthritis body) and the Bupa Health Foundation (a leading charitable organization dedicated to health in Australia). The website consists of a public access area as well as a member resource. This website is publicly accessible; access is not restricted to the participants of this study only.

The information on the website predominantly focused on hip and knee OA. The public access area of the website included the following pages:

1. Treatment and Management Options
2. Fact Sheets
3. Health Care Providers
4. Watch and Listen

The "Treatment and Management Options" and "Fact Sheets" pages provided detailed and evidence-based information covering a variety of topics including conservative treatments, such as weight loss and exercise, as well as surgical treatments such as joint replacement and arthroscopy. The "Health Care Providers" page provided the location of relevant health services within a local area in Australia and the "Watch and Listen" page provided videos that included patient narratives and information about surgery.

Website visitors could choose to complete a validated hip or knee OA risk assessment followed by a visual analog scale for joint pain [27], a cardiovascular risk assessment (as adapted from the Australian Heart Foundation), gastrointestinal (GI) bleeding risk assessment [29], medication and treatment history, and prior consultation with health professionals (in relation to OA). An OA management algorithm based on the answer provided then created a customized management plan for each user.

After creating an account, the members' area reinforced the user's OA management plan and included tailored messages and prompts to encourage users to manage their disease. Weekly, monthly, and biannual assessments allowed members to track pain, weight, treatments and medications, function, and quality of life. The assessments required users to report their pain levels and weight weekly. In addition to these measures, users reported their medication and treatments, completed the Hip disability and Osteoarthritis Outcome Score (HOOS) [30] or the Knee injury and Osteoarthritis Outcome Score (KOOS) [31] and the Quality of Life questionnaires monthly [32]. At 6 months, a Depression, Anxiety and Stress (DASS) [33] assessment and a reevaluation of cardiovascular and gastrointestinal bleeding risk was carried out. This information was processed by an algorithm designed for the website to tailor outputs, such as the management plan, messages that users received based on their needs, and a detailed report that could be discussed with their health care team.

Outcome Measures

Overview

Two reliable and validated tools were used in this study and are described in detail subsequently: the Health Evaluation Impact Questionnaire (heiQ) and the Osteoarthritis Quality Indicator (OAQI).

Health Evaluation Impact Questionnaire

The heiQ is designed to evaluate the effects of health education/self-management programs in terms of health education impact and psychosocial impacts [34]. The 40-item questionnaire is categorized into the following 8 independent domains:

1. Health-directed activity,
2. Positive and active engagement in life,
3. Emotional distress,
4. Self-monitoring and insight,
5. Constructive attitudes and approaches,
6. Skill and technique acquisition,
7. Social integration and support, and
8. Health service navigation.

These various domains are defined to evaluate different aspects of health management including change in behaviors, attitudes and viewpoints, and disease management skills. Imperative to this study, the heiQ evaluated applied knowledge in skill and technique acquisition and self-management in health-directed activity, positive and active engagement in life, and self-monitoring and insights [34].

This instrument identifies motivation of change, adherence with management recommendations, coping, and empowerment.

The heiQ is formatted using a Likert scale with options being strongly disagree, disagree, agree, and strongly agree. For

XSL•FO
RenderX

analysis purposes, these were numerically labeled as 1, 2, 3, and 4, respectively. The score for each domain was calculated as a mean of the contributing questions with higher scores indicating better outcomes for each domain. All questions in the evaluation were asked in the positive with the exception of emotional distress for which lower scores indicated a more favorable outcome. The difference between the means at follow-up compared to baseline provided the change score.

Osteoarthritis Quality Indicator

The OAQI is used to assess the appropriateness of care received by respondents through questions that evaluated patient education and information, clinical assessments, referrals, and treatments [35]. The 17-item tool provides summary pass scores for each item as well as subgroups, which included core treatments versus adjunct treatments and pharmacological treatments versus nonpharmacological treatments (Textbox 1). The OAQI provides insight on adherence to clinical recommendations, aids in identifying barriers in information, and is used in this study to assess change in knowledge acquisition.

Responses to the OAQI were limited to yes, no, and "don't remember," and pass rates were represented as a percentage of positive responses out of all eligible responses (yes and no responses). Pass rates were calculated for each variable for the study sample as a whole and summary pass rates were calculated for each participant for core treatments, adjunct treatments, pharmacological treatments, and nonpharmacological treatment.

To examine the 17 variables within each group and between the intervention and nonparticipating group, the change between follow-up and baseline was tabulated for each participant to indicate negative change, no change, or positive change to provide qualitative change indicators.

Statistical Analysis

Overview

Sample size calculations were based on the primary outcome measure, heiQ, with prior data indicating the difference in the response of matched pairs was normally distributed with a standard deviation of 1.2. To reject the null hypothesis that the response difference was zero with a power of 0.9, 95% significance, and true difference in mean response of 0.6, we needed to recruit 44 participants. However, because this investigation was part of a larger study, recruitment was based on other outcome measures and 300 participants were invited to take part in the study.

Before analysis, normality of data was assessed and participants who had not used the website were identified as nonusers.

Health Evaluation Impact Questionnaire

Changes within the nonusers and the intervention group were analyzed using a paired *t* test comparing follow-up to baseline scores. Differences in change between the groups were compared using an independent *t* test.

Osteoarthritis Quality Indicator

Chi-square tests were used for the evaluation of each individual item within and between the study groups using the qualitative change indicators. The chi-square evaluation for changes within the group compared changes between yes and no responses at baseline and at 12 months. The comparison between groups used the qualitative change indicators (as described previously) in a chi-square evaluation against usage of the website.

Evaluation of the subgroups (core treatments vs adjunct treatments and pharmacological treatments vs nonpharmacological treatments) were carried out using independent *t* tests on each participants change score for each subgroup separately with the use of the website as the independent variable.

XSL•FO
RenderX

Textbox 1. Osteoarthritis Quality Indicator subgroups of domains for core treatments versus adjunct treatments and pharmacological treatments versus nonpharmacological treatments.

Core vs Adjunct Treatments

Core treatments:

- Disease development
- Treatment alternatives
- Self-management
- Lifestyle
- Physical activity
- Referral physical activity
- Weight reduction
- Referral weight reduction

Adjunct treatments:

- Functional assessment
- Walking aid assessment
- Other aids assessment
- Acetaminophen
- Stronger pain killers
- Nonsteroidal anti-inflammatory drugs (NSAIDs)
- Cortisone
- Referral to orthopedic surgeons

Pharmacological vs Nonpharmacological Treatments

Pharmacological treatments:

- Acetaminophen
- Stronger pain killers
- NSAIDs
- Cortisone

Nonpharmacological treatments:

- Disease development
- Treatment alternatives
- Self-management
- Lifestyle
- Physical activity
- Referral physical activity
- Weight reduction
- Referral weight reduction
- Functional assessment
- Walking aid assessment
- Other aids assessment

XSL•FO
RenderX

Results

Participants

A total of 277 eligible participants provided consent and completed baseline measures. As anticipated, a large attrition of volunteers took place and only 195 participants (70.4%) completed the 12-month follow-up questionnaires. Demographic information collected (age, gender, body mass index [BMI], affected joint) showed no significant difference between all collected questionnaires (N=277 including noncompleters), the intervention group (n=104), or the nonusers group (n=91) at baseline (Table 1).

Table 1. Demographic characteristics of participants at baseline (N=277).

Characteristics	Baseline N=277	Noncompleters n=82	Nonusers n=91	My Joint Pain users n=104	*P*
Age (years), mean (SD)	61.0 (8.6)	61.7 (8.5)	60.9 (9.1)	60.5 (8.3)	.64
Female, n (%)	212 (76.5)	60 (73)	73 (80)	79 (76.0)	.54
BMI (kg/m^2), mean (SD)	31.0 (6.8)	30.4 (6.8)	32.5 (10.0)	30.4 (6.8)	.08
Joint, n (%) [a]					.20
Knee	137 (49.4)	45 (55)	38 (42)	54 (52.9)	
Hip	43 (15.5)	7 (9)	16 (18)	20 (19.2)	
Both	92 (33.2)	29 (35)	34 (37)	29 (27.9)	

[a] Total percentage may not add up as it was possible to receive a risk of OA without indicating a joint.

Health Evaluation Impact Questionnaire

Within the intervention group, all domains examined for the heiQ except emotional distress, constructive attitudes and approaches, and health service navigation showed significant improvements from baseline to follow-up. Although differences in baseline mean scores between the 2 groups were minimal (difference in means=0.02) and no significant change was observed in the users, a significant deterioration in emotional distress was observed in the nonusers over the 12-month period (mean difference from baseline to follow-up=0.161, *P*=.008) (Table 2).

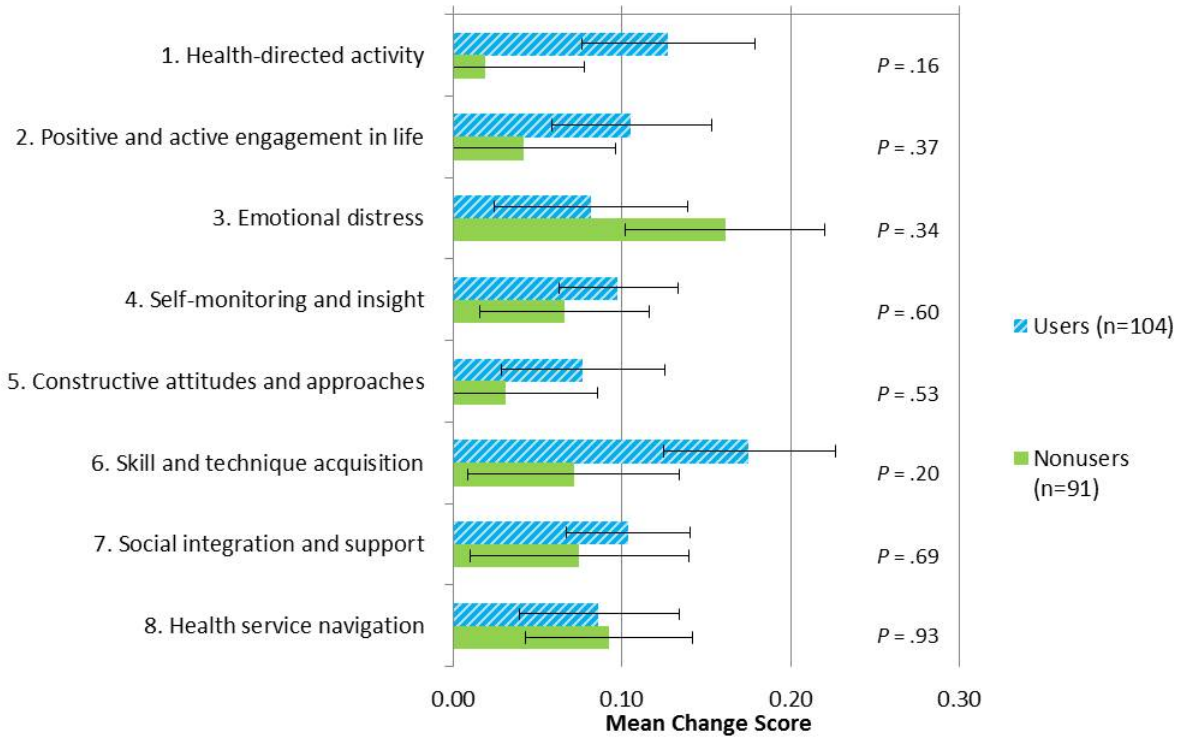
All domains, except health service navigation, showed a clear positive trend in change scores (the difference between follow-up and baseline) in the intervention group relative to nonusers, although the differences in change between the groups were not statistically significant (Figure 1). Mean change scores were calculated as the follow-up score minus the baseline score within each domain and within each group. Positive change scores indicated improvement. *P* values were the result of independent *t* tests comparing the change between groups. All questions in the evaluation were asked in the positive with the exception of emotional distress for which lower scores indicated a more favorable outcome. The improvements appeared particularly pronounced in health-directed activity (*P*=.16), emotional distress (*P*=.34), and skill and technique acquisition (*P*=.20), but none met statistical significance.

Table 2. Mean within-group change between 12-month follow-up and baseline for each Health Evaluation Impact Questionnaire domain for users and nonusers of the My Joint Pain website.

heiQ Domains	Users (n=104)			Nonusers (n=91)		
	Baseline score, mean (SD)	Mean difference (95% CI)	*P*	Baseline score, mean (SD)	Mean difference (95% CI)	*P*
1. Health-directed activity	2.78 (0.71)	0.13 (0.03, 0.23)	.02	2.93 (0.68)	0.02 (–0.10, 0.14)	.74
2. Positive and active engagement in life	2.97 (0.56)	0.11 (0.01, 0.20)	.03	3.10 (0.58)	0.04 (–0.07, 0.15)	.44
3. Emotional distress	2.58 (0.63)	0.08 (–0.03, 0.20)	.16	2.60 (0.82)	0.16 (0.04, 0.28)	.008
4. Self-monitoring and insight	3.05 (0.35)	0.10 (0.03, 0.17)	.01	3.12 (0.43)	0.07 (–0.03, 0.17)	.19
5. Constructive attitudes and approaches	2.99 (0.50)	0.08 (–0.02, 0.17)	.12	3.06 (0.57)	0.03 (–0.08, 0.14)	.58
6. Skill and technique acquisition	2.69 (0.48)	0.18 (0.07, 0.28)	.001	2.79 (0.49)	0.07 (–0.05, 0.20)	.26
7. Social integration and support	2.66 (0.50)	0.10 (0.03, 0.18)	.006	2.64 (0.60)	0.08 (–0.05, 0.20)	.25
8. Health service navigation	2.84 (0.52)	0.09 (–0.01, 0.18)	.07	2.98 (0.59)	0.09 (–0.01, 0.19)	.07

XSL•FO
RenderX

Figure 1. Mean change scores (standard error) between users and nonusers of the My Joint Pain website over a 12-month period for each Health Evaluation Impact Questionnaire domain.



Osteoarthritis Quality Indicator

Pre-post analysis within the intervention group showed significant improvements in self-management (change score=15.2%, *P*=.03), lifestyle (change score=16.2%, *P*=.02), and physical activity (change score=10.8%, *P*=.04) following 12 months' exposure to the website. No significant changes were observed within the nonusers group during this period of time. Compared to nonusers, the intervention group showed significant improvements in self-management (change scores 15.2% vs 1.7%, *P*=.001) and weight reduction (change scores 2.5% vs –6.3%, *P*=.03) measured on the OAQI after 12 months (Table 3).

Summary change scores for treatment categories were also obtained. Between the intervention group and nonusers, there were minimal differences that did not reach statistical significance in core treatments (change scores 10.3% vs 4.7%, *P*=.17), adjunct treatments (change scores 3.1% vs 1.2%, *P*=.54), pharmacological treatments (change scores –1.7% vs –2.4%, *P*=.35), and nonpharmacological treatments (change scores 9.5% vs 4.0%, *P*=.29).

XSL•FO
RenderX

Table 3. Outcomes of the Osteoarthritis Quality Indicator at baseline and 12 months for users and nonusers of the My Joint Pain website.

OAQI items	Users (n=104)			Nonusers (n=91)			*P*
	Baseline, eligible (pass rate)[a]	12-month follow-up, eligible (pass rate)[a]	*P*	Baseline, eligible (pass rate)[a]	12-month follow-up, eligible (pass rate)[a]	*P*	
1. Disease development	99 (56)	95 (68)	.65	81 (53)	81 (59)	.43	.51
2. Treatment alternatives	99 (62)	100 (65)	.62	88 (52)	86 (56)	.64	.80
3. Self-management	99 (53)	99 (68)	.03	88 (47)	87 (48)	.82	.001
4. Lifestyle	99 (53)	99 (69)	.02	89 (48)	90 (53)	.50	.36
5. Physical activity	103 (77)	104 (88)	.04	90 (79)	91 (86)	.23	.79
6. Referral physical activity	104 (51)	103 (62)	.11	90 (53)	89 (57)	.59	.37
7. Weight reduction	74 (69)	70 (71)	.74	64 (83)	64 (77)	.38	.03
8. Referral weight reduction	68 (18)	66 (24)	.35	64 (25)	57 (39)	.11	.95
9. Functional assessment	89 (47)	76 (50)	.72	76 (45)	76 (51)	.42	.40
10. Walking aid assessment	68 (26)	65 (38)	.14	64 (31)	59 (36)	.61	.22
11. Other aids assessment	55 (15)	55 (18)	.61	58 (17)	51 (27)	.20	.80
12. Pain assessment	103 (64)	100 (67)	.66	91 (62)	88 (53)	.27	.37
13. Acetaminophen	104 (76)	103 (81)	.42	91 (79)	89 (79)	.94	.71
14. Stronger pain killers	94 (62)	92 (67)	.42	80 (59)	75 (52)	.40	.62
15. NSAIDs	83 (78)	83 (76)	.71	64 (77)	55 (80)	.65	.41
16. Cortisone	84 (49)	81 (42)	.38	73 (55)	59 (47)	.40	.06
17. Referral to orthopedic surgeons	73 (53)	81 (58)	.57	66 (52)	59 (53)	.91	.29

[a] Number of eligible persons calculated as total number of yes and no answers and who did not provide not applicable or “do not remember” answers. Pass rates are the proportion of yes answers over the number of eligible.

Discussion

Principal Findings and Comparison

The aim of this study was to evaluate the impact of the My Joint Pain website on self-management and quality of care. Our findings indicated that the resource did not improve all aspects of heiQ or OAQI, but highlighted benefits that included improvements in health-directed activity, positive and active engagement in life, self-monitoring and insights, skill and technique acquisition, and social integration within the heiQ. Improvements in self-management, lifestyle, physical activity, and weight reduction were also observed in the OAQI.

The aging of the population and the growing number of people living with chronic conditions is causing increasing dependence on an overstretched health care system. This burden on the health care system coupled with inadequate attention in the clinical context to OA emphasizes the need for self-management strategies and resources for patients with OA [36]. With current clinical practice in OA being suboptimal and self-management programs being ineffective, identifying alternate self-management resources are important [22,37].

Self-management has been increasingly encouraged and supported in various chronic conditions including diabetes, arthritis, and asthma [38]. Unlike traditional patient education in which patients are provided with disease information, self-management education additionally provides techniques that help patients cope and act with their disease including during periods of more active disease [39]. Increasingly, Web-based self-management information and tools have become available with significant efficacy in various diseases including rheumatoid arthritis, fibromyalgia, and juvenile idiopathic arthritis [40-46]. Evidence-based credible tools for OA are needed and the development of the My Joint Pain website was designed to fill this gap by providing tailored, evidence-based information and tools to improve the self-management of the disease and the quality of the clinical consultation between a patient and members of their health care team.

This study hypothesized that the increased and more tailored information received by the participants in the intervention group would empower them to create a collaborative care plan with their health care team, to input their values in treatment choices, and improve self-management. Our hypothesis was partially supported by the results that showed significant improvements in the quality of care patients received as assessed by the OAQI. Specifically, we saw significant improvements in information received about conservative options, such as self-management, lifestyle, and physical activity, as well as

XSL•FO
RenderX

advice regarding weight reduction. These improvements in conservative care knowledge may help to address the existing gap in current clinical practice where management is directed in the first instance to analgesics and then surgery [47]. Furthermore, the improvement in these outcomes could be a result of greater patient confidence in conservative domains that did not require input from a health care professional and where participants could take action independently.

Promising results were also observed when evaluating the impact of the My Joint Pain website with the heiQ. Significant differences in self-monitoring and insights, skill and technique acquisition, and social integration and support were evident within the users of My Joint Pain. Health-directed activity, skill and technique acquisition, and emotional distress were significantly improved for users of the My Joint Pain website and they were the most positively impacted compared to nonusers, although statistical significance was not reached in the differences of change between the groups.

Interestingly, in the heiQ, emotional distress deteriorated in the nonusers but did not change in the intervention group over the 12-month period. This finding contradicted the expectation that improvements would be observed in all domains within the intervention group and none in the nonusers; particularly so in emotional distress as users of the My Joint Pain website had directed encouragement to seek help if biannual check-ups indicated emotional impacts. However, emotional well-being is known to worsen with the progression of OA and it is possible that the website provided a level of maintenance to users to keep their emotional health from deteriorating as observed in the nonusers [48].

The lack of change in the constructive attitude and approaches (indicates how participants view the impact of their disease), and health service navigation domains highlight a gap in the My Joint Pain website. Although there is extensive information on treatments and facts about the disease, there is a lack of information about strategies and techniques to shift ones viewpoint about the impact of OA. In regards to health service navigation, the website does provide a tool to locate nearby health care resources. However, this tool is not easily accessible, well integrated, and not located on the main tool bar, which might account for the observed outcome.

Addressing Barriers to Self-Management

The gaps highlighted from this study relate to a key barrier of self-management in OA: patient's perception. Cuperus et al [49] highlighted the role of patient's perception toward OA, its manageability, the available resources, and their own information needs as both a barrier and a facilitator to self-management. This intervention did not directly aim to modify patient's perception toward self-management and modification of this gap would require the resource to address the concept of self-management. Alternatively, as suggested by Cuperus et al, encouragement and direction toward the resource from health care professionals would be useful [49].

Other established barriers in self-management include limited resources (financial and transportation), low levels of encouragement and support, and a lack of tailored strategies [50]. The My Joint Pain website addresses these barriers by providing a freely available and easily accessible resource as well as using a variety of instruments to tailor recommendations and encouragement for users. The use of tailored features is also identified as a key pillar for the success of Web-based interventions along with other interactive features including self-management, as addressed in My Joint Pain [51]. The additional pillars that were not addressed include social support and contacts with intervention could be addressed by the inclusion of a forum with other consumers and experts.

Similarly, other studies have addressed the inclusion of these pillars to varying degrees in other Web-based arthritic disease interventions, including self-management education, telerehabilitation, and, most commonly, physical activity. Within Internet-based interventions, self-management was investigated in several studies targeted at rheumatoid arthritis, juvenile idiopathic arthritis, fibromyalgia, and OA [44-46]. These studies showed improvements in quality of life and self-rated global health scores. Studies that included more active elements of an intervention, such as personalized phone calls, forums, and Web conferences, also showed positive improvements [45,46,52,53]. Although these interventions were beneficial, the Internet-based interventions were restricted to participants of the study as opposed to the free, open access provided by the My Joint Pain resource; hence, the results are not directly transferable.

Limitations

There are several limitations of this study. There was a large attrition of volunteers with almost one-third lost to follow-up. The reduction in sample size prevented further stratified analysis to evaluate outcomes in varying age groups or genders. Furthermore, information was not collected about disease severity and specific usage data, which prevented the study from correlating changes to specific content or frequency of use of the website. This information would have aided more effective redevelopment of the website. Several other studies have overcome the challenge of identifying usage information by utilizing a website that is password-secured for all the available content [44,52,54]. However, in this study, having a publicly available website prevented log-ins being created solely for participants. Thus, it is possible that the observations may have been attenuated because the intervention group consisted of participants who used the website minimally.

Another challenge a publicly available website as the intervention produced was the inability to structure the study as a randomized controlled trial. Participants were free to use the website, as were the public, at their own convenience. This challenge could have resulted in the observed improvements being driven by a self-administered "attention effect" due to the participants virtue for improvement, in which consumers may have had a strong desire to get better and were active in seeking ways to manage their disease as opposed to being solely driven by the intervention. Inversely, participants in the nonusers group could be unmotivated people who chose not to use the website and hence showed no improvements. Additionally, recruitment was via the Web and the online risk assessment using a self-reported screening tool potentially raises concerns of diagnostic misclassification. Furthermore, the use of the 2

XSL•FO
RenderX

self-reported outcome measures also resulted in a large number of analyses being carried out, which was not corrected for. This paper reports the raw outcomes of the instruments used (Multimedia Appendix 1).

A perceived barrier to the use of a Web-based resource to target older consumers is that they may not be well-versed with Internet usage. However, research has shown that 62% of respondents aged 55 years and older in Australia regularly use the Internet and have both the skills and equipment required to access health information online [55]. Although it is expected that Internet usage among older Australians will reach full saturation in several years, the redevelopment of the website still has to be conscious of the target audience in terms of language and design. However, despite full saturation being expected in several years, other existing means of self-management education, such as booklets, should be employed in conjunction with or parallel to the website to cater to those not using the Internet.

The study outcomes are positive and support the use of a Web-based platform for the dissemination of information and tools. The results of this study will help create a feedback summary for the refinement of the My Joint Pain website to help meet the needs of people at risk of OA or who have already developed this disease. We are now able to identify the gaps and strengths of the intervention that can influence redevelopment. Ideally, the education of patients through this website will improve joint decision making in the clinical consult and result in a higher treatment satisfaction for participants.

Conclusion

An overburdened health care system, an aging population, and increasing Internet usage create the ideal construct for the development of Web-based health care resources. My Joint Pain users showed improvements in several aspects of quality of care received and the result can direct redevelopment to address limitations within the current layout and content of the website. Future studies should be carried out to evaluate its effectiveness on these domains and the website continuously reevaluated to meet the needs of the OA community.

Acknowledgments

This study was funded by an ARC discovery grant (#DP130104407). This website was developed in collaboration with Arthritis Australia in partnership with the Bupa Health Foundation. Dr Hunter is funded by an NHMRC Practitioner Fellowship and Dr Bennell by an NHMRC Principal Research Fellowship. Arthritis Australia received unencumbered funding for the My Joint Pain website from the Bupa Health Foundation. AstraZeneca and Pfizer provided additional unencumbered funding support. Google and Anchor Hosting provided unencumbered in-kind support.

Authors' Contributions

All authors were involved in collecting data, reviewing the literature, and drafting the paper or revising it critically for important intellectual content, and all authors approved the final version to be published.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Raw data from 0 and 12 month OA HUB evaluation.

[XLSX File (Microsoft Excel File), 266KB - jmir_v17i7e167_app1.xlsx]

References

1. Guccione AA, Felson DT, Anderson JJ, Anthony JM, Zhang Y, Wilson PW, et al. The effects of specific medical conditions on the functional limitations of elders in the Framingham Study. Am J Public Health 1994 Mar;84(3):351-358. [Medline: 8129049]
2. Henderson JV, Harrison CM, Britt HC, Bayram CF, Miller GC. Prevalence, causes, severity, impact, and management of chronic pain in Australian general practice patients. Pain Med 2013 Sep;14(9):1346-1361. [doi: 10.1111/pme.12195] [Medline: 23855874]
3. Knox SA, Harrison CM, Britt HC, Henderson JV. Estimating prevalence of common chronic morbidities in Australia. Med J Aust 2008 Jul 21;189(2):66-70. [Medline: 18637769]
4. Centers for Disease Control and Prevention. Prevalence of disabilities and associated health conditions among adults--United States, 1999. MMWR Morb Mortal Wkly Rep 2001 Feb 23;50(7):120-125 [FREE Full text] [Medline: 11393491]
5. Lawrence RC, Felson DT, Helmick CG, Arnold LM, Choi H, Deyo RA, National Arthritis Data Workgroup. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part II. Arthritis Rheum 2008 Jan;58(1):26-35 [FREE Full text] [doi: 10.1002/art.23176] [Medline: 18163497]

XSL•FO
RenderX

6. Jordan KM, Arden NK, Doherty M, Bannwarth B, Bijlsma JWJ, Dieppe P, Standing Committee for International Clinical Studies Including Therapeutic Trials ESCISIT. EULAR Recommendations 2003: an evidence based approach to the management of knee osteoarthritis: Report of a Task Force of the Standing Committee for International Clinical Studies Including Therapeutic Trials (ESCISIT). Ann Rheum Dis 2003 Dec;62(12):1145-1155 [FREE Full text] [Medline: 14644851]
7. Zhang W, Doherty M, Arden N, Bannwarth B, Bijlsma J, Gunther K, EULAR Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). EULAR evidence based recommendations for the management of hip osteoarthritis: report of a task force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). Ann Rheum Dis 2005 May;64(5):669-681 [FREE Full text] [doi: 10.1136/ard.2004.028886] [Medline: 15471891]
8. Richmond J, Hunter D, Irrgang J, Jones MH, Levy B, Marx R, American Academy of Orthopaedic Surgeons. Treatment of osteoarthritis of the knee (nonarthroplasty). J Am Acad Orthop Surg 2009 Sep;17(9):591-600 [FREE Full text] [Medline: 19726743]
9. National Health and Medical Research Council. Guideline for the Non-Surgical Management of Hip and Knee Osteoarthritis. South Melbourne: The Royal Australian College of General Practitioners; 2009 Jul. URL: http://www.racgp.org.au/download/documents/Guidelines/Musculoskeletal/racgp_oa_guideline.pdf [accessed 2015-06-19] [WebCite Cache ID 6ZOd2W7RF]
10. Hochberg MC, Altman RD, April KT, Benkhalti M, Guyatt G, McGowan J, American College of Rheumatology. American College of Rheumatology 2012 recommendations for the use of nonpharmacologic and pharmacologic therapies in osteoarthritis of the hand, hip, and knee. Arthritis Care Res (Hoboken) 2012 Apr;64(4):465-474. [Medline: 22563589]
11. McAlindon TE, Bannuru RR, Sullivan MC, Arden NK, Berenbaum F, Bierma-Zeinstra SM, et al. OARSI guidelines for the non-surgical management of knee osteoarthritis. Osteoarthritis Cartilage 2014 Mar;22(3):363-388 [FREE Full text] [doi: 10.1016/j.joca.2014.01.003] [Medline: 24462672]
12. DeHaan MN, Guzman J, Bayley MT, Bell MJ. Knee osteoarthritis clinical practice guidelines -- how are we doing? J Rheumatol 2007 Oct;34(10):2099-2105. [Medline: 17722223]
13. Dhawan A, Mather RC, Karas V, Ellman MB, Young BB, Bach BR, et al. An epidemiologic analysis of clinical practice guidelines for non-arthroplasty treatment of osteoarthritis of the knee. Arthroscopy 2014 Jan;30(1):65-71. [doi: 10.1016/j.arthro.2013.09.002] [Medline: 24290788]
14. Porcheret M, Jordan K, Jinks C, Croft P, Primary Care Rheumatology Society. Primary care treatment of knee pain--a survey in older adults. Rheumatology (Oxford) 2007 Nov;46(11):1694-1700 [FREE Full text] [doi: 10.1093/rheumatology/kem232] [Medline: 17938135]
15. Brand CA, Harrison C, Tropea J, Hinman RS, Britt H, Bennell K. Management of osteoarthritis in general practice in Australia. Arthritis Care Res (Hoboken) 2014 Apr;66(4):551-558. [doi: 10.1002/acr.22197] [Medline: 24127305]
16. Mitchell HL, Hurley MV. Management of chronic knee pain: a survey of patient preferences and treatment received. BMC Musculoskelet Disord 2008;9:123 [FREE Full text] [doi: 10.1186/1471-2474-9-123] [Medline: 18801169]
17. Arthritis Australia. The Ignored Majority: The Voice of Arthritis 2011. 2011. URL: http://www.arthritisaustralia.com.au/images/stories/documents/reports/2011_updates/the%20voice%20of%20arthritis%202011.pdf [accessed 2015-02-17] [WebCite Cache ID 6WOv5HTZ3]
18. Ong LM, de Haes JC, Hoos AM, Lammes FB. Doctor-patient communication: a review of the literature. Soc Sci Med 1995 Apr;40(7):903-918. [Medline: 7792630]
19. Barlow J. How to use education as an intervention in osteoarthritis. Best Pract Res Clin Rheumatol 2001 Oct;15(4):545-558. [doi: 10.1053/berh.2001.0172] [Medline: 11567538]
20. Cabana MD, Rand CS, Powe NR, Wu AW, Wilson MH, Abboud PA, et al. Why don't physicians follow clinical practice guidelines? A framework for improvement. JAMA 1999 Oct 20;282(15):1458-1465. [Medline: 10535437]
21. Holman H, Lorig K. Patients as partners in managing chronic disease. Partnership is a prerequisite for effective and efficient health care. BMJ 2000 Feb 26;320(7234):526-527 [FREE Full text] [Medline: 10688539]
22. Kroon FP, van der Burg LR, Buchbinder R, Osborne RH, Johnston RV, Pitt V. Self-management education programmes for osteoarthritis. Cochrane Database Syst Rev 2014;1:CD008963. [doi: 10.1002/14651858.CD008963.pub2] [Medline: 24425500]
23. Ricciardi L, Mostashari F, Murphy J, Daniel JG, Siminerio EP. A national action plan to support consumer engagement via e-health. Health Aff (Millwood) 2013 Feb;32(2):376-384. [doi: 10.1377/hlthaff.2012.1216] [Medline: 23381531]
24. Autralian Bureau of Statistics. 2011 Dec 15. Household Use of Information Technology, Australia, 2010-11 URL: http://www.abs.gov.au/ausstats/abs@.nsf/0/4E4D83E02F39FC32CA25796600152BF4?opendocument [accessed 2015-02-17] [WebCite Cache ID 6WOvSp4Bs]
25. Glasgow RE, McKay HG, Piette JD, Reynolds KD. The RE-AIM framework for evaluating interventions: what can it tell us about approaches to chronic illness management? Patient Educ Couns 2001 Aug;44(2):119-127. [Medline: 11479052]
26. My Joint Pain. URL: https://myjointpain.org.au/ [accessed 2015-06-19] [WebCite Cache ID 6ZOfg79Ep]
27. Kwoh CK. A screening tool for knee osteoarthritis. In: Arthritis Rheum. 63 Suppl. 2011 Presented at: Association of Rheumatology Health Professionals Annual Scientific Meeting; November 4-9, 2011; Chicago, IL p. 10 URL: http://www.blackwellpublishing.com/acrmeeting/abstract.asp?MeetingID=781&id=96819

XSL•FO
RenderX

28. International Patient Decision Aids Standards (IPDAS) Collaboration. IPDAS 2005: Criteria for Judging the Quality of Patient Decision Aids. URL: http://www.ipdas.ohri.ca/ipdas_checklist.pdf [accessed 2015-02-17] [WebCite Cache ID 6WOsscgt5]
29. Singh G, Ramey DR, Triadafilopoulus G, Brown BW, Balise RR. GI score: A simple self-assessment instrument to quantify the risk of serious NSAID-related GI complications in RA and OA. 1998 Presented at: American College of Rheumatology; November 1998; San Diego, CA p. S75-S75.
30. Nilsdotter AK, Lohmander LS, Klässbo M, Roos EM. Hip disability and osteoarthritis outcome score (HOOS)--validity and responsiveness in total hip replacement. BMC Musculoskelet Disord 2003 May 30;4:10 [FREE Full text] [doi: 10.1186/1471-2474-4-10] [Medline: 12777182]
31. Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynnon BD. Knee Injury and Osteoarthritis Outcome Score (KOOS)--development of a self-administered outcome measure. J Orthop Sports Phys Ther 1998 Aug;28(2):88-96. [doi: 10.2519/jospt.1998.28.2.88] [Medline: 9699158]
32. Nakamura A, Osonoi T, Terauchi Y. Relationship between urinary sodium excretion and pioglitazone-induced edema. J Diabetes Investig 2010 Oct 19;1(5):208-211 [FREE Full text] [doi: 10.1111/j.2040-1124.2010.00046.x] [Medline: 24843434]
33. Lovibond PF, Lovibond SH. The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. Behav Res Ther 1995 Mar;33(3):335-343. [Medline: 7726811]
34. Osborne RH, Elsworth GR, Whitfield K. The Health Education Impact Questionnaire (heiQ): an outcomes and evaluation measure for patient education and self-management interventions for people with chronic conditions. Patient Educ Couns 2007 May;66(2):192-201. [doi: 10.1016/j.pec.2006.12.002] [Medline: 17320338]
35. Østerås N, Garratt A, Grotle M, Natvig B, Kjeken I, Kvien TK, et al. Patient-reported quality of care for osteoarthritis: development and testing of the osteoarthritis quality indicator questionnaire. Arthritis Care Res (Hoboken) 2013 Jul;65(7):1043-1051. [doi: 10.1002/acr.21976] [Medline: 23401461]
36. 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]
37. Runciman WB, Hunt TD, Hannaford NA, Hibbert PD, Westbrook JI, Coiera EW, et al. CareTrack: assessing the appropriateness of health care delivery in Australia. Med J Aust 2012 Jul 16;197(2):100-105. [Medline: 22794056]
38. Newman S, Steed L, Mulligan K. Self-management interventions for chronic illness. Lancet 2004;364(9444):1523-1537. [doi: 10.1016/S0140-6736(04)17277-2] [Medline: 15500899]
39. 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]
40. Lorig K, Ritter PL, Laurent DD, Plant K, Green M, Jernigan VBB, 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]
41. 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]
42. Farvolden P, Denisoff E, Selby P, Bagby RM, Rudy L. Usage and longitudinal effectiveness of a Web-based self-help cognitive behavioral therapy program for panic disorder. J Med Internet Res 2005;7(1):e7 [FREE Full text] [doi: 10.2196/jmir.7.1.e7] [Medline: 15829479]
43. Joseph CLM, Peterson E, Havstad S, Johnson CC, Hoerauf S, Stringer S, Asthma in Adolescents Research Team. A web-based, tailored asthma management program for urban African-American high school students. Am J Respir Crit Care Med 2007 May 1;175(9):888-895 [FREE Full text] [doi: 10.1164/rccm.200608-1244OC] [Medline: 17290041]
44. 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]
45. Shigaki CL, Smarr KL, Siva C, Ge B, Musser D, Johnson R. RAHelp: an online intervention for individuals with rheumatoid arthritis. Arthritis Care Res (Hoboken) 2013 Oct;65(10):1573-1581. [doi: 10.1002/acr.22042] [Medline: 23666599]
46. Stinson JN, McGrath PJ, Hodnett ED, Feldman BM, Duffy CM, Huber AM, et al. An internet-based self-management program with telephone support for adolescents with arthritis: a pilot randomized controlled trial. J Rheumatol 2010 Sep;37(9):1944-1952. [doi: 10.3899/jrheum.091327] [Medline: 20595280]
47. Hunter DJ, Lo GH. The management of osteoarthritis: an overview and call to appropriate conservative treatment. Rheum Dis Clin North Am 2008 Aug;34(3):689-712. [doi: 10.1016/j.rdc.2008.05.008] [Medline: 18687278]
48. Riddle DL, Kong X, Fitzgerald GK. Psychological health impact on 2-year changes in pain and function in persons with knee pain: data from the Osteoarthritis Initiative. Osteoarthritis Cartilage 2011 Sep;19(9):1095-1101 [FREE Full text] [doi: 10.1016/j.joca.2011.06.003] [Medline: 21723400]
49. Cuperus N, Smink AJ, Bierma-Zeinstra SM, Dekker J, Schers HJ, de Boer F, et al. Patient reported barriers and facilitators to using a self-management booklet for hip and knee osteoarthritis in primary care: results of a qualitative interview study. BMC Fam Pract 2013;14:181 [FREE Full text] [doi: 10.1186/1471-2296-14-181] [Medline: 24289303]
50. Bair MJ, Matthias MS, Nyland KA, Huffman MA, Stubbs DL, Kroenke K, et al. Barriers and facilitators to chronic pain self-management: a qualitative study of primary care patients with comorbid musculoskeletal pain and depression. Pain Med 2009 Oct;10(7):1280-1290 [FREE Full text] [doi: 10.1111/j.1526-4637.2009.00707.x] [Medline: 19818038]

XSL•FO
RenderX

51. Morrison LG, Yardley L, Powell J, Michie S. What design features are used in effective e-health interventions? A review using techniques from Critical Interpretive Synthesis. Telemed J E Health 2012 Mar;18(2):137-144. [doi: 10.1089/tmj.2011.0062] [Medline: 22381060]
52. van den Berg MH, Ronday HK, Peeters AJ, Voogt-van der Harst EM, Munneke M, Breedveld FC, et al. Engagement and satisfaction with an Internet-based physical activity intervention in patients with rheumatoid arthritis. Rheumatology (Oxford) 2007 Mar;46(3):545-552 [FREE Full text] [doi: 10.1093/rheumatology/kel341] [Medline: 17043050]
53. Russell TG, Buttrum P, Wootton R, Jull GA. Internet-based outpatient telerehabilitation for patients following total knee arthroplasty: a randomized controlled trial. J Bone Joint Surg Am 2011 Jan 19;93(2):113-120. [doi: 10.2106/JBJS.I.01375] [Medline: 21248209]
54. Bossen D, Veenhof C, Van Beek KE, Spreeuwenberg PM, Dekker J, De Bakker DH. Effectiveness of a web-based physical activity intervention in patients with knee and/or hip osteoarthritis: randomized controlled trial. J Med Internet Res 2013;15(11):e257 [FREE Full text] [doi: 10.2196/jmir.2662] [Medline: 24269911]
55. Burns P, Jones SC, Iverson D, Caputi P. Riding the wave or paddling in the shallows? Understanding older Australians' use of the internet. Health Promot J Austr 2012 Aug;23(2):145-148. [Medline: 23088477]

Abbreviations

heiQ: Health Evaluation Impact Questionnaire
HON: Health on Net
OA: osteoarthritis
OAQI: Osteoarthritis Quality Indicator

Edited by G Eysenbach; submitted 24.02.15; peer-reviewed by N Østerås, C Brand; comments to author 19.03.15; revised version received 20.04.15; accepted 05.05.15; published 07.07.15.

Please cite as:
Umapathy H, Bennell K, Dickson C, Dobson F, Fransen M, Jones G, Hunter DJ
The Web-Based Osteoarthritis Management Resource My Joint Pain Improves Quality of Care: A Quasi-Experimental Study
J Med Internet Res 2015;17(7):e167
URL: http://www.jmir.org/2015/7/e167/
doi:10.2196/jmir.4376
PMID:26154022

©Hema Umapathy, Kim Bennell, Chris Dickson, Fiona Dobson, Marlene Fransen, Graeme Jones, David J Hunter. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 07.07.2015. 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

Why Breast Cancer Risk by the Numbers Is Not Enough: Evaluation of a Decision Aid in Multi-Ethnic, Low-Numerate Women

Rita Kukafka[1], MA, DrPH; Haeseung Yi[2], MPA; Tong Xiao[3], MPH; Parijatham Thomas[4], MD; Alejandra Aguirre[5], MPH; Cindy Smalletz[6], MA; Raven David[7], BA; Katherine Crew[8], MD, MS

[1]Columbia University, College of Physicians and Surgeons, Biomedical Informatics, Mailman School of Public Health, Sociomedical Sciences, New York, NY, United States

[2]Columbia University, Teachers College, Health and Behavior Studies, New York, NY, United States

[3]Mailman School of Public Health, Columbia University, Epidemiology, New York, NY, United States

[4]MD Anderson Cancer Center, Department of Clinical Cancer Prevention, Houston, TX, United States

[5]Columbia University, Community Engagement Resource, Irving Institute for Clinical & Translational Research, New York, NY, United States

[6]Columbia University, Center for New Media Teaching and Learning, New York, NY, United States

[7]Columbia University, Biomedical Informatics, New York, NY, United States

[8]Columbia University, College of Physicians and Surgeons, Medicine, Herbert Irving Comprehensive Cancer Center, New York, NY, United States

Corresponding Author:
Rita Kukafka, MA, DrPH
Columbia University, College of Physicians and Surgeons
Biomedical Informatics
Mailman School of Public Health, Sociomedical Sciences
622 West 168th Street, PH-20, Room 314
New York, NY,
United States
Phone: 1 212 580 5560
Fax: 1 212 580 5560
Email: rk326@cumc.columbia.edu

Abstract

Background: Breast cancer risk assessment including genetic testing can be used to classify people into different risk groups with screening and preventive interventions tailored to the needs of each group, yet the implementation of risk-stratified breast cancer prevention in primary care settings is complex.

Objective: To address barriers to breast cancer risk assessment, risk communication, and prevention strategies in primary care settings, we developed a Web-based decision aid, RealRisks, that aims to improve preference-based decision-making for breast cancer prevention, particularly in low-numerate women.

Methods: RealRisks incorporates experience-based dynamic interfaces to communicate risk aimed at reducing inaccurate risk perceptions, with modules on breast cancer risk, genetic testing, and chemoprevention that are tailored. To begin, participants learn about risk by interacting with two games of experience-based risk interfaces, demonstrating average 5-year and lifetime breast cancer risk. We conducted four focus groups in English-speaking women (age ≥18 years), a questionnaire completed before and after interacting with the decision aid, and a semistructured group discussion. We employed a mixed-methods approach to assess accuracy of perceived breast cancer risk and acceptability of RealRisks. The qualitative analysis of the semistructured discussions assessed understanding of risk, risk models, and risk appropriate prevention strategies.

Results: Among 34 participants, mean age was 53.4 years, 62% (21/34) were Hispanic, and 41% (14/34) demonstrated low numeracy. According to the Gail breast cancer risk assessment tool (BCRAT), the mean 5-year and lifetime breast cancer risk were 1.11% (SD 0.77) and 7.46% (SD 2.87), respectively. After interacting with RealRisks, the difference in perceived and estimated breast cancer risk according to BCRAT improved for 5-year risk (P=.008). In the qualitative analysis, we identified potential barriers to adopting risk-appropriate breast cancer prevention strategies, including uncertainty about breast cancer risk and risk models, distrust toward the health care system, and perception that risk assessment to pre-screen women for eligibility for genetic testing may be viewed as rationing access to care.

XSL•FO
RenderX

Conclusions: In a multi-ethnic population, we demonstrated a significant improvement in accuracy of perceived breast cancer risk after exposure to RealRisks. However, we identified potential barriers that suggest that accurate risk perceptions will not suffice as the sole basis to support informed decision making and the acceptance of risk-appropriate prevention strategies. Findings will inform the iterative design of the RealRisks decision aid.

(J Med Internet Res 2015;17(7):e165) doi:10.2196/jmir.4028

KEYWORDS

breast cancer; decision making; risk communication; consumer health informatics; genetic testing; decision aid; risk stratified screening

Introduction

Breast cancer confers significant morbidity and mortality on women in the United States, and the primary prevention of this disease is a major public health issue. In 2014, an estimated 232,670 women in the United States will be diagnosed with breast cancer and 40,000 women will die from this disease [1]. Known breast cancer risk factors include age, family history, benign breast disease, reproductive history, and lifestyle factors, such as alcohol intake and obesity [2]. Genetic determinants, such as BRCA1 and BRCA2 mutations, confer the greatest impact on breast cancer risk with a 40-60% lifetime risk [3]. The Gail breast cancer risk assessment tool (BCRAT) is the most commonly used model in the United States and provides an individual's absolute 5-year and lifetime risk of invasive breast cancer compared to the general population [4]. Breast cancer risk assessment, including genetic testing for hereditary breast cancer, is underutilized in the United States [5]. Many women may be unaware of their risk status due to our inability to adequately screen them in the primary care setting. Using a risk-stratified approach, breast cancer screening and preventive options could be tailored to an individual's risk profile to maximize benefits and minimize harms. Barriers to adopting risk-appropriate screening and prevention include inaccurate risk perceptions, inadequate time for counseling, insufficient knowledge about risk-reducing strategies, and a number of potential ethical and social issues [5-8].

Women from racial/ethnic minorities are less likely to seek breast cancer preventive care [9,10], contributing to higher rates of late stage diagnosis and poorer clinical outcomes in these populations compared to non-Hispanic whites [11-13]. Low numeracy (ie, the effective use of quantitative information to guide health behavior and make health decisions) affects up to 93 million Americans and constrains counseling about cancer risk and prevention strategies [14,15]. We previously reported that low-numerate patients were more likely than high-numerate patients to overestimate their risk [15]. Most research to date that has focused on explaining risks to patients in narratives, numbers, or graphs reveals that all three forms can produce reasoning biases that complicate risk communication [15-17]. People overweigh rare events when they read described probabilities but assign them lower weight when they experience probabilities through an activity such as drawing cards from a deck [18,19]. When participants used an experience-based dynamic interface to interpret risk in our previous study, the differences in risk perceptions associated with low numeracy were reduced [20].

The purpose of this study is to conduct focus groups to inform the iterative design of RealRisks, a patient-centered decision aid for communicating breast cancer risk, reducing inaccurate risk perceptions, and providing preference-based decision support for risk management. When integrated into clinical workflow, RealRisks will identify high-risk women, present them with a unique experience-based dynamic interface to communicate risk, and facilitate communication between patients and clinicians about the risks and benefits of appropriate preventive strategies. We targeted a multi-ethnic population of women from New York City with a high proportion of low numeracy.

Methods

Recruitment

In June 2013, we conducted four focus groups among English-speaking women, age ≥18 years, recruited from Northern Manhattan in New York, New York. Women who participated in a community database through the Community Engagement Core Resource of the Irving Institute for Clinical and Translational Research were contacted via email or telephone. Each focus group consisted of 7-9 women. The study was approved by the Institutional Review Board at Columbia University Medical Center, and all participants provided written informed consent. A total of 34 women that reside in the Washington Heights/Inwood community participated.

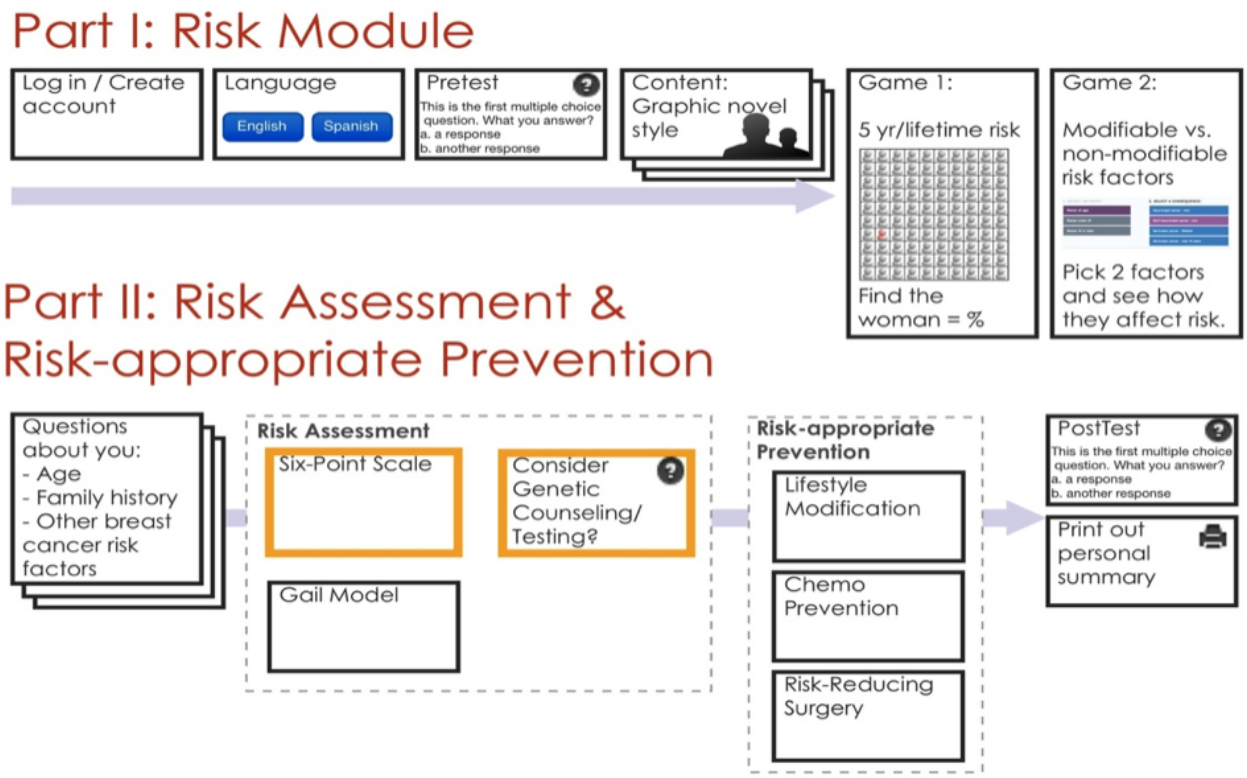
Description of the RealRisks Decision Aid

RealRisks models patient-provider dialogue and incorporates experience-based dynamic interfaces to communicate numeric and probabilistic concepts that are central to risk communication (Figure 1). The narrative is based on a fictitious character Rose, who has a family history of breast cancer and visits her doctor for a routine check-up. We segmented the narrative into the following modules: (1) risk (what is risk, what are breast cancer risk factors), (2) genetic testing (hereditary breast cancer, inherited mutations), and (3) chemoprevention (anti-estrogens, risks/benefits). Embedded within the narrative of RealRisks are two games of experience-based risk interfaces, based upon our previous work [21]. The games demonstrate absolute 5-year and lifetime breast cancer risk for an average 50-year-old woman using a pictograph with 100 clickable women. Players are instructed to click until they "find" a woman with breast cancer. Players continue to click (eg, sample from the population of 100 women) to better learn the meaning of a given pre-set probability (ie, 12 out of 100 women or 12%). Our data suggest that this interactive experience-based format for representing

XSL•FO
RenderX

risk improves accuracy of risk perception in a low-numerate population [21].

Figure 1. Schema of the RealRisks decision aid.



Conducting the Focus Groups

A skilled facilitator (ANA) led the focus groups using detailed guides (available from the authors upon request). The discussion guide included questions on breast cancer risk factors, BRCA genetic testing, and discussing breast cancer risk and/or genetic testing with a provider. The sessions lasted about 90 minutes and were audio recorded. For the first 15 minutes, women participated in a discussion about their experiences with breast cancer and how they understood breast cancer risk. For the next 30 minutes, the participants were allowed to view and interact with the RealRisks decision aid on a laptop and listen to an audio recording of the narrative. The last 30 minutes involved a semistructured group discussion to obtain feedback on the acceptability and usefulness of the DA. Specifically, we were interested in learning: (1) Do users accept the decision aid for learning about breast cancer risk and genetic testing?, (2) Can users easily navigate and use the decision aid?, (3) Does the decision aid effectively increase users' confidence and participation in the decision-making process?, and (4) Does the decision aid increase knowledge, understanding of breast cancer risk, and risk management options in the context of an individual's risk profile?

Prior to starting the focus group discussions, a self-administered questionnaire including information about demographics, numeracy [22], Internet access, sources of information, and breast cancer risk factors was completed at baseline. Perceived breast cancer risk using a validated measure [23] was assessed before and after exposure to RealRisks and evaluation of the decision aid on a 7-point Likert scale was administered post intervention.

Quantitative Analysis

Descriptive statistics were generated to document participant baseline characteristics and frequencies of positive and negative attitudes about the RealRisks decision aid. Perceived breast cancer risk was assessed by asking "What is your best guess about your percent chance of developing breast cancer during the next 5 years?" and "during your lifetime?" on a scale from 0% to 100% [22]. The Gail BCRAT was used to estimate absolute 5-year and lifetime invasive breast cancer risk [4]. Accuracy of perceived breast cancer risk was defined as within ±5% of estimated lifetime risk according to the BCRAT. Paired *t* test and McNemar's test were used to compare within-individual changes in accuracy of perceived breast cancer risk before and after interacting with RealRisks.

Qualitative Analysis

For the qualitative analysis, 2 investigators (HSY and TX) independently read the transcript from the first completed focus group to develop the initial codes and coding template. We identified meaningful segments within the responses and assigned codes using an editing style analysis [24]. Discrepancies in coding were negotiated at weekly research meetings. HSY and TX independently read and coded the remaining three focus group transcripts, applying the coding template, which was iteratively modified as the analysis proceeded. We grouped codes into general themes and discussed the themes among the entire team of investigators. The team collectively selected the themes and representative quotes presented in this paper. Atlas.ti 7.0 software was used to facilitate qualitative data management and analysis. All transcripts were uploaded into the software to enable

XSL•FO
RenderX

investigators to do coding, build the codebook, and group the codes into themes. A final comparison of coding across all interviews yielded 62.3%-94.5% agreement. Thematic saturation was determined based on consensus of the coders that new thematic categories were no longer arising from the final focus group transcript.

Results

Participant Characteristics

The baseline characteristics of the 34 focus group participants are summarized below (Table 1). The majority (62%, 21/34) were Hispanic, and mean age was 53 years (range 35-75). A total of 41% (14/34) met criteria for low numeracy, defined as a score of 0-5 (range 0-9) [23], and 65% (22/34) demonstrated poor knowledge of breast cancer risk factors, defined as a score of 0-9 (range 0-18). Everyone had access to the Internet, including 89% (30/34) who used the Internet at least weekly. In terms of breast cancer risk factors, 24% (8/34) of women had a first-degree family history of breast cancer and 13% (4/34) had a prior benign breast biopsy. According to the BCRAT (excluding 3 women with a history of breast cancer), mean absolute 5-year and lifetime risk were 1.11% (range 0.2%-4.3%) and 7.46% (range 2.8%-14.6%), respectively, and 10% (3/34) of women met high-risk criteria for breast cancer (≥1.67% 5-year risk).

Table 1. Baseline characteristics of focus group participants (N=34), New York City (2013).

Characteristics of focus group participants	
Age (years), mean (SD)	53.4 (10.2)
Race/ethnicity, n (%)	
Non-Hispanic white	2 (5.9)
Non-Hispanic black	8 (23.5)
Hispanic	21 (61.8)
Asian	1 (2.9)
Other	2 (5.9)
Low numeracy[a], n (%)	14 (41.2)
Poor knowledge of breast cancer risk factors[b], n (%)	22 (64.7)
First-degree family history of breast cancer, n (%)	8 (23.5)
Prior benign breast biopsy, n (%)	4 (12.9)
High risk for breast cancer[c], n (%)	3 (9.7)
5-year breast cancer risk[d], mean (SD)	1.11 (0.77)
Lifetime breast cancer risk[d], mean (SD)	7.46 (2.87)

[a]Numeracy score ranges from 0-9. Low numeracy defined as a score of 0-5 [23].

[b]Score of knowledge of breast cancer risk factors ranges from 0-18, with poor knowledge defined as a score of 0-9.

[c]According to the BCRAT, high risk is defined as 5-year invasive breast cancer risk ≥1.67%.

[d]Excluding 3 women with a prior history of breast cancer.

Quantitative Analysis

Perceived 5-year and lifetime breast cancer risk ranged from 0-100%. After interacting with RealRisks, the difference in perceived and estimated breast cancer risk according to the BCRAT significantly improved for 5-year risk (P=.008), but not for lifetime risk (P=.20) (Table 2). Accuracy of perceived breast cancer risk improved from 52% to 70% (P=.10). Even in the subgroup of women with low numeracy, accurate risk perceptions improved from 46% to 70%. In particular, 80% (4/5) women who overestimated their lifetime breast cancer risk by more than 30% had accurate risk perceptions after exposure to RealRisks.

Participants' impressions about the RealRisks decision aid on a 7-point Likert scale are shown in Table 3. Over 75% found the decision aid easy to use and felt that it increased their knowledge about breast cancer, genetic testing, and chemoprevention. Over 87% (29/33) would recommend RealRisks to a friend.

XSL•FO
RenderX

Table 2. Accuracy of breast cancer risk perceptions before and after interacting with the RealRisks decision aid among focus group participants (N=34), New York City (2013).

Breast cancer risk perception	Before RealRisks	After RealRisks	*P* value
Perceived 5-year breast cancer risk (%), mean (SD)	10.4 (22.4)	5.3 (12.1)	.008[a]
Perceived lifetime breast cancer risk (%), mean (SD)	13.1 (26.1)	9.6 (13.7)	.20[a]
Accurate perceived breast cancer risk [b] , n (%)	15 (51.7)	19 (70.4)	.10[c]
High numeracy	10 (55.6)	12 (70.6)	
Low numeracy	5 (45.5)	7 (70.0)	

[a]*P* value based upon paired *t* test.

[b]Accurate perceived breast cancer risk defined as within ±5% of estimated lifetime breast cancer risk according to the BCRAT.

[c]*P* value based upon McNemar's test.

Table 3. Evaluation of the RealRisks decision aid on a 7-point Likert scale among focus group participants, New York City (2013).

	Frequency, n (%)		
	Disagree (1-2)	Neutral (3-5)	Agree (6-7)
1. RealRisks is useful	1 (3.0)	9 (27.2)	23 (69.7)
2. Most would learn how to use quickly	0 (0)	13 (39.4)	20 (60.6)
3. Easy to use	0 (0)	4 (12.1)	29 (87.9)
4. Increased knowledge of breast cancer	0 (0)	3 (9.1)	30 (90.9)
5. Increased knowledge of genetic testing	0 (0)	6 (18.2)	27 (81.8)
6. Increased knowledge of chemoprevention	0 (0)	8 (24.2)	25 (75.8)
7. Helped to understand breast cancer risk	0 (0)	11 (33.3)	22 (66.7)
8. Helped to understand lifetime breast cancer risk	1 (3.0)	6 (18.2)	26 (78.8)
9. Helped to understand modifiable risk factors	1 (3.0)	11 (33.3)	21 (63.6)
10. I can relate to Rose	4 (12.1)	9 (27.3)	20 (60.6)
11. Will help to discuss genetic testing with doctor	3 (9.1)	7 (21.2)	23 (69.7)
12. Will help to discuss chemoprevention with doctor	4 (12.1)	8 (24.2)	21 (63.6)
13. More confident about decision making about genetic testing	0 (0)	8 (24.2)	25 (75.8)
14. Less worried about getting breast cancer	2 (6.1)	12 (36.4)	19 (57.6)
15. Women have a choice about getting genetic testing	1 (3.0)	4 (12.1)	28 (84.8)
16. Would recommend RealRisks to a friend	0 (0)	4 (12.1)	29 (87.9)

Qualitative Analysis

Overview

In spite of the improvements in accuracy of perceived breast cancer risk after exposure to RealRisks, three major factors emerged as potential barriers to adoption of risk-appropriate breast cancer prevention strategies (Textbox 1): (1) uncertainty about breast cancer risk and risk models, (2) distrust toward the health care system, and (3) perception that risk assessment to pre-screen women for eligibly for genetic testing may be a proxy for rationing access to care.

XSL•FO
RenderX

Textbox 1. Potential barriers to adopting risk-appropriate breast cancer prevention strategies among focus group participants, New York City (2013).

Uncertainty about breast cancer risk and risk models:

"As far as constants are concerned, if we were going to do scientific research, the only thing constant is change. I think we are all exposed to things that we are not in control of. So, I think it's optimistic of us to look at our lineage, to look at people in our family who have had it [breast cancer]. But I think because of external effects, it is hard for us to determine what percentage we might be more susceptible to."

"I have no idea what the percentage is. No one in my family had breast cancer. They had cysts in their breasts. I have no idea. Nobody smokes in my home. But walking around here, I'm exposed to it."

"You can't do catch-up. If it's in you, you are going to get it [breast cancer]."

"I think the risk [for breast cancer] for everybody is 100%...because anyone can get it."

"I have 50% chance. My mother died from breast cancer and I had a breast surgery. My daughter had a big lump in her chest at 12. My father had prostate cancer. It's a horrible death. I still hear my mother screaming."

"I would like to know if I have the gene mutation. So for me to say that being my mom had it [breast cancer] and take a pill and prevent it and not knowing if I have the gene doesn't make sense to me."

Distrust toward the health care system:

"No, because we don't have the knowledge and don't ask, and the doctor won't up and give it to you."

"If you ask your doctor in the clinic, he's going to look at you how do you know that? With that look. The condescending attitude. Where did you hear about that?"

"I would look for several doctors. I wouldn't believe in one doctor's result I swear because they are all human. The way medicine is going today, it's not as humanistic as it was. They are not talking to you in the face. They are staring at a bloody computer screen. Dentists are the same. The whole way of practicing medicine today is so highly technical. I would have to seek out a second, third, or maybe even a fourth doctor before I make a decision. I don't base my decision strictly on what the doctor says. They've been wrong at times."

"Unfortunately we all need to be very proactive and never accept no for an answer. Simply because one or two people said no, they are not going to do the [genetic] test...I think you have to keep going and if it's something deep down in your heart that is someone should get the test, then they should receive that test."

"Medical establishment has certain guidelines and recently there is some conflict about mammograms. There is a group, it's like a renegade group and say you don't need a mammogram every year, you just need it every 3 or 4 years. Don't listen to them, because they are rebels. These groups also said you don't need to do it in your 40s, you just need to do it in your 50s or 60s. These groups always pop up."

Access to care:

"Why would the health insurance pay for the chemoprevention pill but not for the genetic test to see if you have the gene? I would have an issue with my insurance."

"The [genetic] test that Angelina Jolie took was $3000, so right now how can everybody find out if they can't afford it...or if they don't have insurance?"

"My feeling is that the only deterrent [to genetic testing] would be the cost. If it's just a blood test, not invasive, what could possibly stop me? For me, personally, it would be the $3000. That's a biggie...it's the only thing that would be stopping me."

"The ones with money can get all this testing done. For people that don't have money, they are out there."

Uncertainty About Breast Cancer Risk and Risk Models

A consistent theme in our analysis was uncertainty among participants that all factors were accounted for in determining their breast cancer risk. The most commonly expressed concern was the potential for the risk assessment tool to miss something that was unknown and outside their control.

Similar to several prior studies on risk communication, participants referred to quantitative risk with many misconceptions. Having a close family member with cancer appeared to influence a participant's feeling of risk. Women that relayed their experience with a close friend or family member often used the proverbial "50% chance" to express the idea that anyone can get breast cancer.

Participants viewed any risk, even a low risk, with a degree of uncertainty and therefore not acceptable as the basis for their eligibility to get genetic testing. The perceptions among participants was the "test might uncover the unexpected". According to one participant who stated her breast cancer risk is 2%: "I would still want to know. There's always that possibility. There is always that chance and I may not be in the high-risk group and that might be the key to finding out if I carry it [genetic mutation]". These comments highlight the need for educating patients about when genetic test results are informative or uninformative for estimating cancer risk.

Distrust Toward the Health Care System

None of the participants recalled discussing breast cancer risk with their physician. Notably, they expressed their own lack of knowledge as a reason for not initiating a discussion and doubted that their physician would engage. Participants expressed a degree of responsibility to seek out the information they needed to inform their decisions and the need to be proactive because providers won't "give it up".

The recent debate about mammographic screening guidelines was discussed in the context of suspicion and distrust. Participants speculated about possible ulterior motives, such as the lack of funding for diseases that concerned women and these

XSL•FO
RenderX

perceptions led to reduced credibility of the information they received. These passages highlight how source credibility (providers, health care system in general) may extend perceived uncertainty in breast cancer risk assessment tools.

Access to Care

In responding to barriers to genetic testing, not having money was discussed as primary. The financial constraints expressed by women led to many concerns that lack of money or insurance would limit their access. Removing the financial barrier, many of the women expressed the desire to receive genetic testing. As stated by one participant: “If it’s free, I’ll go every 6 months”. Stated by another women: “The ones with money can get all this testing done. For people that don’t have money, they are out there”. Only one woman voiced concern that a woman might regret a decision to “cut their breasts off” based upon their genetic testing results, stating that genetic testing is far too soon to be necessary because “some people will panic”. These comments highlight our participants’ view that use of risk assessment to pre-screen women for eligibly for genetic testing may be perceived as service rationing.

Discussion

Principal Findings

In this community of largely Hispanic women from Upper Manhattan, we demonstrated a significant improvement in accuracy of breast cancer risk perception after interacting with the experience-based dynamic interfaces to communicate risk embedded in the RealRisks decision aid. This improvement was found in high numeracy women and also in low numeracy women who systematically have been shown to overestimate their breast cancer risk [25,26].

However, consistent with previous research [27,28], our qualitative analysis showed that women do not associate their disease risk with access to health services—for example, those who perceived their risk to be low still want to be tested for breast cancer susceptibility genes. Thus, more accurate risk perceptions may not suffice as the sole basis to support clinical decision making for breast cancer risk management. We identified three major themes as potential barriers to adoption of risk-appropriate breast cancer prevention strategies. These findings have several implications for communicating risk management options based upon breast cancer risk prediction models.

First, there is a great deal of uncertainty regarding breast cancer risk and the utility of the risk prediction models. There may be misconceptions about testing for breast cancer susceptibility genes—for example, as previously shown, some may incorrectly believe that the BRCA test can diagnose cancer [29]. Prior studies suggest that individuals often reject their personal risk estimates. In the context of colon cancer, one study found that of the participants who correctly remembered their personalized risk of getting colon cancer, only half actually accepted it as valid [30]. With respect to breast cancer, approximately 20% of women did not believe that their personalized risk number was accurate [31]. Women in our study expressed a level of uncertainty that all factors used to determine their breast cancer risk were accounted for. They appeared to justify their perceived risk as better or worse based on their understanding of less well-defined breast cancer risk factors that are not included in the BCRAT model. For example, participants who regularly engage in physical activity may have perceived their personal breast cancer risk as low, whereas participants mentioned being exposed to air pollutants in an urban environment leads to a higher than average breast cancer risk. These so called rational adjustments to risk estimates were similarly shown in a study to evaluate the Alzheimer’s disease risk perceptions of individuals who accurately recall their genetics-based Alzheimer disease risk assessment [32]. While not explicitly discussed, the utterances of the women in our focus groups recognized that risk models are imperfect tools. Drawing an analogy to Taleb’s “black swan” logic [33], the inability to predict outliers (black swans) implies the inability to predict those that lie outside the realm of regular expectations. Women frequently gave examples of what we are not in control of and unexpected events (eg, the 35-year-old woman diagnosed with breast cancer without clear risk factors).

Prominent in our analysis was the perception that health care itself was the source of distrust, which contributes to the uncertainty of the breast cancer risk prediction models. It is well understood that one of the most important determinants of responses to risk information is likely to be whether the information source is perceived to be credible and trustworthy. Eagley, Wood, and Chaiken have identified two types of communicator bias that message recipients might infer [34]. The first is knowledge bias, which refers to a belief that the communicator’s knowledge about truth is adequate. In the context of risk assessment models, this might refer to the perception that the tool itself can calculate probabilities to express uncertainty (eg, are the important factors considered and is the science about known risk factors adequate?). The second is reporting bias. In the context of our study, this might refer to a belief that the provider and health care system more generally is not distorting information in order to promote a particular view. Our participant’s statement that it is the “renegade group” saying you no longer need a mammogram every year expresses the concern that recommendations about screening mammography are distorted and should not be trusted. While our study was not designed to explore the mechanisms contributing to sources of distrust and how such perceptions might influence risk models, our findings build on prior work suggesting that cancer risk assessment models can be viewed with uncertainty [35,36] and distrust [27,30,31] and that source credibility is a key determinant [35,36].

Another novel finding was the perception of using risk models to guide breast cancer risk management options as a barrier to access to care. Some of the women in our focus groups viewed not being eligible for genetic testing as service rationing. As one woman stated “only the ones with money can get all this testing done”. The well-known Angelina Jolie effect resulted in demand for BRCA testing to almost double [37]. Although a survey carried out in the United States found that although 75% of Americans were aware of Angelina Jolie’s double mastectomy, fewer than 10% of respondents had the information necessary to accurately interpret her risk of developing cancer

XSL•FO
RenderX

relative to a woman unaffected by the BRCA gene mutation [38]. Awareness of the Angelina Jolie story was not associated with improved understanding. The women in our study estimated the cost of genetic testing to be about US $3000, which raised the question "right now, how can everybody find out if they can't afford it?". In order to avoid undermining wider trust in the health care system, effective communication strategies are needed to ensure that those designated as low-risk understand and trust that the rationale behind tailoring prevention regimens is not about reducing or withholding services, but rather it is to optimize benefits and reduce potential harms. The complexities of this communication could be most problematic for women with low numeracy, as they tend to overestimate their risk. Therefore, active communication and assurances are critical to mitigating any exacerbation of existing disparities.

While underutilized in the United States, breast cancer risk assessment including genetic testing can be used to classify people into different risk groups with screening and preventive interventions tailored to the needs of each group [5]. By risk-stratifying the population, screening and management options could be applied differentially to each population stratum with potentially more efficient allocation of resources. However, even if improved health outcomes are achieved, the implementation of risk-stratified breast cancer prevention programs has proven to be complex [39-41]. Currently, there is a scarcity of empirical evidence about patient acceptance and their preferences with regard to risk-appropriate cancer screening and prevention [41].

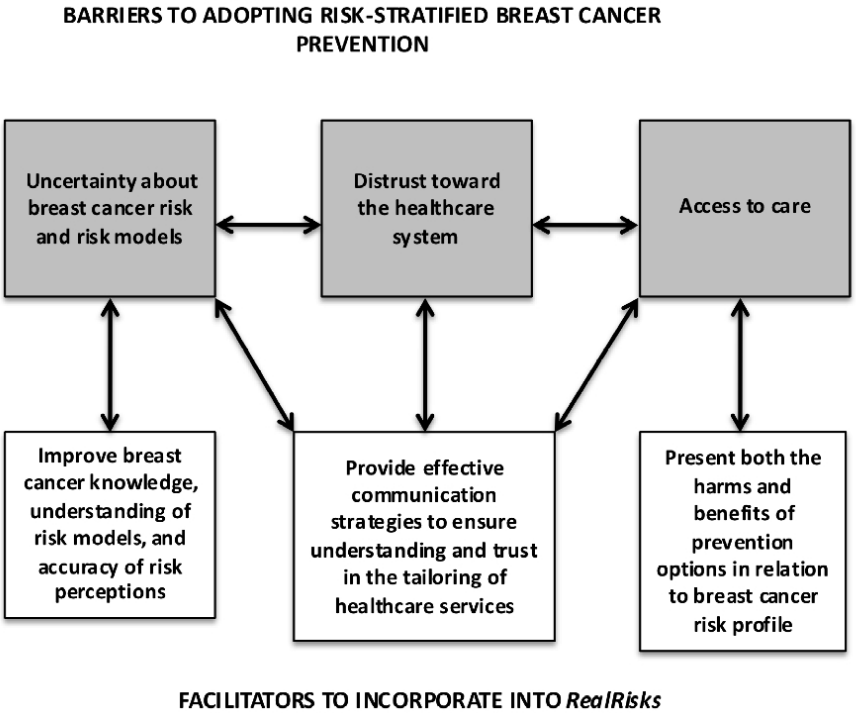
Limitations

Our urban-dwelling female participants, who were largely minorities with a high proportion of low numeracy, may not be representative of the general population in other geographic areas. However, given that this is an understudied and underserved population, this is also a unique strength of our study. There may have also been selection bias, where women with a greater interest in expressing their opinions about breast cancer may have been more likely to participate, which affects the generalizability of our results.

Conclusions

We conducted this study to inform the iterative design of the RealRisks decision aid. While the experience-based interface resulted in improved accuracy of breast cancer risk perceptions, our qualitative findings identified additional barriers to risk-based health care delivery, which need to be addressed (Figure 2). For example, to address the theme of distrust toward the health care system, RealRisks will now incorporate dialogue to explain that genetic testing has few or no benefits for women who do not have a family history that is associated with increased risk for BRCA1 or BRCA2 mutations. The experience-based interface will be extended to include how taking chemoprevention pills might impact their personalized risk, so they can learn by interacting with the game that benefit is seen only among high-risk women and risks outweigh benefits for women below a specified risk threshold. Moreover, we will emphasize using both dialogue and games that women across all risk strata have preventive options. The game will allow women to learn about screening and lifestyle choices. We consider this to be particularly salient as controversy over the potential harms of population-based mammographic screening due to overdiagnosis continues to escalate [42,43]. Future studies are needed to determine how these iterations to RealRisks are received, and more generally, whether decision aids, such as RealRisks, can improve accuracy of breast cancer risk perceptions, informed decision making, and acceptance of risk-appropriate prevention strategies.

Figure 2. Schema of barriers and facilitators to the adoption of breast cancer risk assessment and risk-appropriate prevention strategies, which will inform the iterative design and refinement of the RealRisks decision aid.



XSL•FO
RenderX

Acknowledgments

Financial support for this study was provided in part by grants from the National Cancer Institute (Grant Number R01 CA177995-01A1), the National Center for Advancing Translational Sciences (Grant Number UL1 TR000040), and Women at Risk. The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report.

Conflicts of Interest

None declared.

References

1. National Cancer Institute Surveillance, Edipemiology, and End Results Program. SEER Stat Fact Sheets: Breast Cancer URL: http://seer.cancer.gov/statfacts/html/breast.html [accessed 2014-11-13] [WebCite Cache ID 6U2gtgCkM]
2. Singletary SE. Rating the risk factors for breast cancer. Ann Surg 2003 Apr;237(4):474-482. [doi: 10.1097/01.SLA.0000059969.64262.87] [Medline: 12677142]
3. Chen S, Parmigiani G. Meta-analysis of BRCA1 and BRCA2 penetrance. J Clin Oncol 2007 Apr 10;25(11):1329-1333 [FREE Full text] [doi: 10.1200/JCO.2006.09.1066] [Medline: 17416853]
4. Costantino JP, Gail MH, Pee D, Anderson S, Redmond CK, Benichou J, et al. Validation studies for models projecting the risk of invasive and total breast cancer incidence. J Natl Cancer Inst 1999 Sep 15;91(18):1541-1548 [FREE Full text] [Medline: 10491430]
5. Bellcross CA, Leadbetter S, Alford SH, Peipins LA. Prevalence and healthcare actions of women in a large health system with a family history meeting the 2005 USPSTF recommendation for BRCA genetic counseling referral. Cancer Epidemiol Biomarkers Prev 2013 Apr;22(4):728-735 [FREE Full text] [doi: 10.1158/1055-9965.EPI-12-1280] [Medline: 23371291]
6. Pace LE, Keating NL. A systematic assessment of benefits and risks to guide breast cancer screening decisions. JAMA 2014 Apr 2;311(13):1327-1335. [doi: 10.1001/jama.2014.1398] [Medline: 24691608]
7. Rosenbaum L. Invisible risks, emotional choices--mammography and medical decision making. N Engl J Med 2014 Oct 16;371(16):1549-1552. [doi: 10.1056/NEJMms1409003] [Medline: 25317876]
8. Biller-Andorno N, Jüni P. Abolishing mammography screening programs? A view from the Swiss Medical Board. N Engl J Med 2014 May 22;370(21):1965-1967. [doi: 10.1056/NEJMp1401875] [Medline: 24738641]
9. Kaplan CP, Haas JS, Pérez-Stable EJ, Gregorich SE, Somkin C, Des JG, et al. Breast cancer risk reduction options: awareness, discussion, and use among women from four ethnic groups. Cancer Epidemiol Biomarkers Prev 2006 Jan;15(1):162-166 [FREE Full text] [doi: 10.1158/1055-9965.EPI-04-0758] [Medline: 16434605]
10. Jacobson JS, Grann VR, Hershman D, Troxel AB, Li H, Neugut AI. Breast biopsy and race/ethnicity among women without breast cancer. Cancer Detect Prev 2006;30(2):129-133. [doi: 10.1016/j.cdp.2006.02.002] [Medline: 16621329]
11. Armstrong K, Micco E, Carney A, Stopfer J, Putt M. Racial differences in the use of BRCA1/2 testing among women with a family history of breast or ovarian cancer. JAMA 2005 Apr 13;293(14):1729-1736. [doi: 10.1001/jama.293.14.1729] [Medline: 15827311]
12. Hall MJ, Olopade OI. Disparities in genetic testing: thinking outside the BRCA box. J Clin Oncol 2006 May 10;24(14):2197-2203. [doi: 10.1200/JCO.2006.05.5889] [Medline: 16682739]
13. Ward EM, Smith RA. Integrating tools for breast cancer risk assessment, risk reduction, and early detection. Cancer Epidemiol Biomarkers Prev 2010 Oct;19(10):2428-2429 [FREE Full text] [doi: 10.1158/1055-9965.EPI-10-0835] [Medline: 20929880]
14. Lipkus IM, Peters E, Kimmick G, Liotcheva V, Marcom P. Breast cancer patients' treatment expectations after exposure to the decision aid program adjuvant online: the influence of numeracy. Med Decis Making 2010;30(4):464-473 [FREE Full text] [doi: 10.1177/0272989X09360371] [Medline: 20160070]
15. Ancker JS, Kaufman D. Rethinking health numeracy: a multidisciplinary literature review. J Am Med Inform Assoc 2007;14(6):713-721 [FREE Full text] [doi: 10.1197/jamia.M2464] [Medline: 17712082]
16. McNeil BJ, Pauker SG, Sox HC, Tversky A. On the elicitation of preferences for alternative therapies. N Engl J Med 1982 May 27;306(21):1259-1262. [doi: 10.1056/NEJM198205273062103] [Medline: 7070445]
17. Fagerlin A, Wang C, Ubel PA. Reducing the influence of anecdotal reasoning on people's health care decisions: is a picture worth a thousand statistics? Med Decis Making 2005;25(4):398-405. [doi: 10.1177/0272989X05278931] [Medline: 16061891]
18. Weber EU, Shafir S, Blais A. Predicting risk sensitivity in humans and lower animals: risk as variance or coefficient of variation. Psychol Rev 2004 Apr;111(2):430-445. [doi: 10.1037/0033-295X.111.2.430] [Medline: 15065916]
19. Hertwig R, Barron G, Weber EU, Erev I. Decisions from experience and the effect of rare events in risky choice. Psychol Sci 2004 Aug;15(8):534-539. [doi: 10.1111/j.0956-7976.2004.00715.x] [Medline: 15270998]

XSL•FO
RenderX

20. Ancker JS, Weber EU, Kukafka R. Effect of arrangement of stick figures on estimates of proportion in risk graphics. Med Decis Making 2011;31(1):143-150. [doi: 10.1177/0272989X10369006] [Medline: 20671209]
21. Ancker JS, Weber EU, Kukafka R. Effects of game-like interactive graphics on risk perceptions and decisions. Med Decis Making 2011;31(1):130-142. [doi: 10.1177/0272989X10364847] [Medline: 20393103]
22. Lipkus IM, Kuchibhatla M, McBride CM, Bosworth HB, Pollak KI, Siegler IC, et al. Relationships among breast cancer perceived absolute risk, comparative risk, and worries. Cancer Epidemiol Biomarkers Prev 2000 Sep;9(9):973-975 [FREE Full text] [Medline: 11008917]
23. Lipkus IM, Samsa G, Rimer BK. General performance on a numeracy scale among highly educated samples. Med Decis Making 2001;21(1):37-44. [Medline: 11206945]
24. Crabtree BF, Miller WL. Negotiating the coresearcher mandate - service users' experiences of doing collaborative research on mental health. In: Doing qualitative research. Thousand Oaks, CA: Sage Publications; 1999.
25. Black WC, Nease RF, Tosteson AN. Perceptions of breast cancer risk and screening effectiveness in women younger than 50 years of age. J Natl Cancer Inst 1995 May 17;87(10):720-731. [Medline: 7563149]
26. Schwartz LM, Woloshin S, Black WC, Welch HG. The role of numeracy in understanding the benefit of screening mammography. Ann Intern Med 1997 Dec 1;127(11):966-972. [Medline: 9412301]
27. Henneman L, Timmermans DR, Bouwman CM, Cornel MC, Meijers-Heijboer H. 'A low risk is still a risk': exploring women's attitudes towards genetic testing for breast cancer susceptibility in order to target disease prevention. Public Health Genomics 2011;14(4-5):238-247. [doi: 10.1159/000276543] [Medline: 20090298]
28. Bruno M, Digennaro M, Tommasi S, Stea B, Danese T, Schittulli F, et al. Attitude towards genetic testing for breast cancer susceptibility: a comparison of affected and unaffected women. Eur J Cancer Care (Engl) 2010 May;19(3):360-368. [doi: 10.1111/j.1365-2354.2009.01067.x] [Medline: 19912305]
29. Daly MB, Lerman CL, Ross E, Schwartz MD, Sands CB, Masny A. Gail model breast cancer risk components are poor predictors of risk perception and screening behavior. Breast Cancer Res Treat 1996;41(1):59-70. [Medline: 8932877]
30. Weinstein ND, Atwood K, Puleo E, Fletcher R, Colditz G, Emmons KM. Colon cancer: risk perceptions and risk communication. J Health Commun 2004;9(1):53-65. [doi: 10.1080/10810730490271647] [Medline: 14761833]
31. Scherer LD, Ubel PA, McClure J, Greene SM, Alford SH, Holtzman L, et al. Belief in numbers: When and why women disbelieve tailored breast cancer risk statistics. Patient Educ Couns 2013 Aug;92(2):253-259 [FREE Full text] [doi: 10.1016/j.pec.2013.03.016] [Medline: 23623330]
32. Linnenbringer E, Roberts JS, Hiraki S, Cupples LA, Green RC. "I know what you told me, but this is what I think:" perceived risk of Alzheimer disease among individuals who accurately recall their genetics-based risk estimate. Genet Med 2010 Apr;12(4):219-227 [FREE Full text] [doi: 10.1097/GIM.0b013e3181cef9e1] [Medline: 20139767]
33. Taleb N. The Black Swan: The Impact of the Highly Improbable. New York: Random House Trade Paperbacks; 2010.
34. Eagly AH, Wood W, Chaiken S. Causal inferences about communicators and their effect on opinion change. Journal of Personality and Social Psychology 1978;36(4):424-435. [doi: 10.1037/0022-3514.36.4.424]
35. Han PKJ, Klein WMP, Lehman TC, Massett H, Lee SC, Freedman AN. Laypersons' responses to the communication of uncertainty regarding cancer risk estimates. Med Decis Making 2009;29(3):391-403 [FREE Full text] [doi: 10.1177/0272989X08327396] [Medline: 19470720]
36. Johnson BB, Slovic P. Presenting uncertainty in health risk assessment: initial studies of its effects on risk perception and trust. Risk Anal 1995 Aug;15(4):485-494. [Medline: 7480948]
37. Evans DG, Barwell J, Eccles DM, Collins A, Izatt L, Jacobs C, et al. The Angelina Jolie effect: how high celebrity profile can have a major impact on provision of cancer related services. Breast Cancer Res 2014;16(5):442 [FREE Full text] [doi: 10.1186/s13058-014-0442-6] [Medline: 25510853]
38. Borzekowski DLG, Guan Y, Smith KC, Erby LH, Roter DL. The Angelina effect: immediate reach, grasp, and impact of going public. Genet Med 2014 Jul;16(7):516-521. [doi: 10.1038/gim.2013.181] [Medline: 24357847]
39. Hall AE, Chowdhury S, Hallowell N, Pashayan N, Dent T, Pharoah P, et al. Implementing risk-stratified screening for common cancers: a review of potential ethical, legal and social issues. J Public Health (Oxf) 2014 Jun;36(2):285-291 [FREE Full text] [doi: 10.1093/pubmed/fdt078] [Medline: 23986542]
40. Pashayan N, Duffy SW, Chowdhury S, Dent T, Burton H, Neal DE, et al. Polygenic susceptibility to prostate and breast cancer: implications for personalised screening. Br J Cancer 2011 May 10;104(10):1656-1663 [FREE Full text] [doi: 10.1038/bjc.2011.118] [Medline: 21468051]
41. Nelson HD, Pappas M, Zakher B, Mitchell JP, Okinaka-Hu L, Fu R. Risk assessment, genetic counseling, and genetic testing for BRCA-related cancer in women: a systematic review to update the U.S. Preventive Services Task Force recommendation. Ann Intern Med 2014 Feb 18;160(4):255-266. [doi: 10.7326/M13-1684] [Medline: 24366442]
42. Kalager M, Adami H, Bretthauer M, Tamimi RM. Overdiagnosis of invasive breast cancer due to mammography screening: results from the Norwegian screening program. Ann Intern Med 2012 Apr 3;156(7):491-499. [doi: 10.7326/0003-4819-156-7-201204030-00005] [Medline: 22473436]
43. Bleyer A, Welch HG. Effect of three decades of screening mammography on breast-cancer incidence. N Engl J Med 2012 Nov 22;367(21):1998-2005. [doi: 10.1056/NEJMoa1206809] [Medline: 23171096]

XSL•FO
RenderX

Abbreviations

BCRAT: breast cancer risk assessment tool
BRCA: BReast CAncer

Edited by G Eysenbach; submitted 02.12.14; peer-reviewed by M Jimbo; comments to author 14.03.15; revised version received 17.04.15; accepted 27.04.15; published 14.07.15.

Please cite as:
Kukafka R, Yi H, Xiao T, Thomas P, Aguirre A, Smalletz C, David R, Crew K
Why Breast Cancer Risk by the Numbers Is Not Enough: Evaluation of a Decision Aid in Multi-Ethnic, Low-Numerate Women
J Med Internet Res 2015;17(7):e165
URL: http://www.jmir.org/2015/7/e165/
doi:10.2196/jmir.4028
PMID:26175193

©Rita Kukafka, Haeseung Yi, Tong Xiao, Parijatham Thomas, Alejandra Aguirre, Cindy Smalletz, Raven David, Katherine Crew. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 14.07.2015. 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

Improving Pediatric Basic Life Support Performance Through Blended Learning With Web-Based Virtual Patients: Randomized Controlled Trial

Ronny Lehmann[1], MD, MME; Christiane Thiessen[1]; Barbara Frick[1], MD; Hans Martin Bosse[2], MD, MME; Christoph Nikendei[3], MD, MME; Georg Friedrich Hoffmann[1], MD; Burkhard Tönshoff[1], MD; Sören Huwendiek[4], MD, MME

[1]Center for Pediatrics and Adolescent Medicine, Department of General Pediatrics, University Hospital Heidelberg, Heidelberg, Germany

[2]Center for Pediatrics and Adolescent Medicine, Department of General Pediatrics, University Hospital Düsseldorf, Düsseldorf, Germany

[3]Department of General Internal Medicine and Psychosomatics, University Hospital Heidelberg, Heidelberg, Germany

[4]Institute of Medical Education, Department of Assessment and Evaluation, Faculty of Medicine, University of Bern, Bern, Switzerland

Corresponding Author:
Ronny Lehmann, MD, MME
Center for Pediatrics and Adolescent Medicine
Department of General Pediatrics
University Hospital Heidelberg
Im Neuenheimer Feld 430
Heidelberg, 69120
Germany
Phone: 49 62215639360
Fax: 49 6221565618
Email: ronny.lehmann@med.uni-heidelberg.de

Abstract

Background: E-learning and blended learning approaches gain more and more popularity in emergency medicine curricula. So far, little data is available on the impact of such approaches on procedural learning and skill acquisition and their comparison with traditional approaches.

Objective: This study investigated the impact of a blended learning approach, including Web-based virtual patients (VPs) and standard pediatric basic life support (PBLS) training, on procedural knowledge, objective performance, and self-assessment.

Methods: A total of 57 medical students were randomly assigned to an intervention group (n=30) and a control group (n=27). Both groups received paper handouts in preparation of simulation-based PBLS training. The intervention group additionally completed two Web-based VPs with embedded video clips. Measurements were taken at randomization (t0), after the preparation period (t1), and after hands-on training (t2). Clinical decision-making skills and procedural knowledge were assessed at t0 and t1. PBLS performance was scored regarding adherence to the correct algorithm, conformance to temporal demands, and the quality of procedural steps at t1 and t2. Participants' self-assessments were recorded in all three measurements.

Results: Procedural knowledge of the intervention group was significantly superior to that of the control group at t1. At t2, the intervention group showed significantly better adherence to the algorithm and temporal demands, and better procedural quality of PBLS in objective measures than did the control group. These aspects differed between the groups even at t1 (after VPs, prior to practical training). Self-assessments differed significantly only at t1 in favor of the intervention group.

Conclusions: Training with VPs combined with hands-on training improves PBLS performance as judged by objective measures.

(J Med Internet Res 2015;17(7):e162) doi:10.2196/jmir.4141

KEYWORDS

virtual patients; blended learning; simulation; pediatric basic life support; performance

XSL•FO
RenderX

Introduction

Basic life support training, such as for pediatric basic life support (PBLS), is usually simulation-based with the need for evaluating learners' performances [1-4]. Although there is evidence that simulator training is effective to improve basic life support performance, literature comparing various methods of training is scarce [5]. In particular, the instructional design of life support training is increasingly being investigated. Carrero et al assessed the improvement in procedural knowledge acquired by typically used tutor-led, case-based discussions versus the use of noninteractive multimedia presentations—video plus PowerPoint presentation. Both were shown to have equal impact on the level of cognitive skills [6]. Some reports have shown advantages for learning basic life support when using instructional videos [7-9]. Such approaches provide individual preparation and can be easily distributed, save instructors' resources, and allow for more training time in face-to-face sessions.

For promoting clinical reasoning and decision making, virtual patients (VPs) are known for being effective [10]. For the context of acquiring life support skills, VPs integrate features that have been shown to foster both the development of clinical decision making (eg, through interactivity and feedback [11]) and procedural skills (eg, by integration of media [12]). E-learning and blended learning approaches are gaining popularity in emergency medicine curricula [13-16]. Lehmann et al reported recently that VPs combined with skills laboratory training are perceived by both trainees and trainers as an effective approach to train undergraduates in PBLS, leading to an efficient use of training time [17]. A few other reports have already suggested positive effects of VPs and comparable simulators regarding knowledge and procedural skill acquisition used for different kinds of life support courses [18-20].

In this study, we investigated the effect of VPs combined with standard simulation-based PBLS training on the acquisition of clinical decision-making skills and procedural knowledge, objective skill performance, and self-assessment. Our hypotheses were that preparation with VPs would yield (1) superior clinical decision making and procedural knowledge, (2) an objectively better performance of PBLS after the training, and (3) better self-assessment after working with VPs and after exposure to standard training.

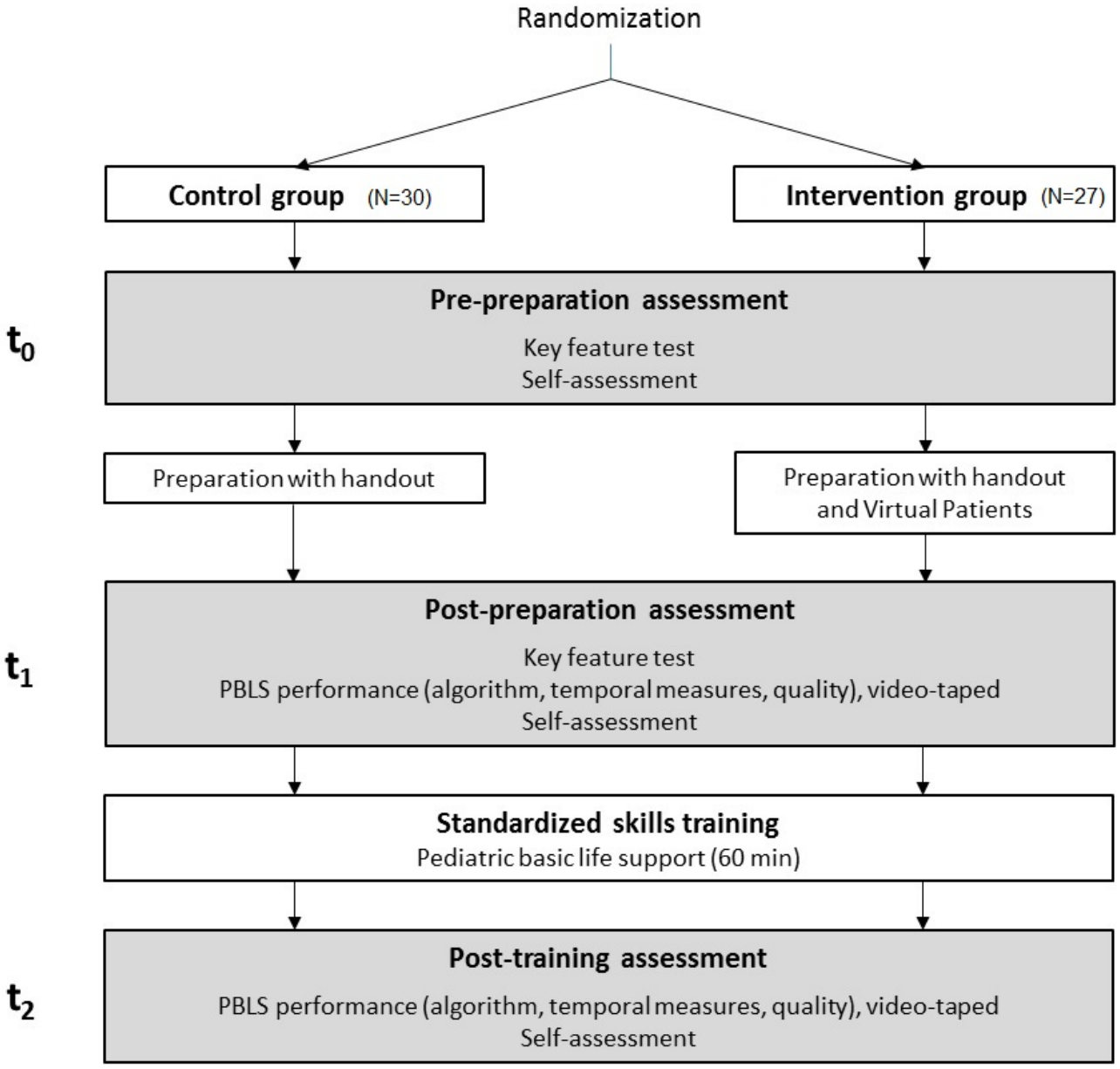
Methods

Study Design

We used a two-group randomized trial design (see Figure 1). All participants were assessed regarding their self-assessment, clinical decision-making skills, and procedural knowledge (key-feature test) about PBLS after randomization to ensure comparability (prepreparation assessment, t_0). PBLS training sessions were conducted 1 to 2 weeks after the preparation assessment. Both groups were requested to prepare themselves a day ahead of the appointed training using handouts we had distributed. In addition, the intervention group (IG) was granted access to VPs as mandatory preparation. After the preparation, on the day of the practical training, self-assessment and procedural knowledge were assessed again to compare the participants' progress (postpreparation assessment, t_1). Subsequently, we videotaped PBLS sequences undertaken by each participant for later scoring of their performances. Both groups then attended standard training on PBLS. Later that day, we again recorded PBLS demonstrations and reevaluated participants' self-assessments after the practical training (posttraining assessment, t_2). The study was conducted in September 2014.

XSL•FO
RenderX

Figure 1. Study design.



Instruments

Overview

All instruments were pilot-tested on video recordings of PBLS demonstrations by student tutors and faculty before implementation, and revisions were made to ensure clarity and content validity. We particularly tested the estimated and calculated temporal scores adapted from international recommendations [21] by recording and analyzing best-practice examples of our faculty.

Basic Data

Participants were asked about their age, sex, and level of qualification in emergency medicine. For subgroup analysis we identified participants who were qualified as paramedics or had some similar training—qualifications that include PBLS training.

Clinical Decision-Making Skills and Procedural Knowledge

We developed a key-feature test according to published guidelines [22] to evaluate the students' procedural knowledge and clinical decision making. This kind of testing was introduced by Page and Bordage specifically to assess clinical decision-making skills [23]. The test contained seven cases with three key features each (see Multimedia Appendix 1). Answers were to be given in "write in" format, which was suggested for decisions regarding the differential diagnosis, therapy, and further management [22]. Questions concerned both clinical decision making (proposed next steps) and procedural knowledge (eg, head positioning or compression depth). Each correct answer was given 1 point, with a maximum of 21 points. The test was reviewed for correctness and clinical relevance by group-blinded senior pediatricians with expertise in PBLS.

XSL•FO
RenderX

Performance: Adherence to Algorithm

Two raters scored the performed algorithm for its correct order. Each step of the sequence was given 2 points if it was done in the correct algorithmic order. It was given 1 point if it had been performed in an incorrect algorithmic order. No points were assigned if the step had not been undertaken at all (see Multimedia Appendix 2). The maximum score was 18.

Performance: Temporal Demands

Concrete temporal recommendations for three procedural steps of the PBLS algorithm are as follows [21].

1. Every rescue breath should take 1.0 to 1.5 s for inspiration plus time for expiration.

2. Assessment of the signs of life and circulation may not take longer than 10 s.

3. Chest compressions should be given at a frequency of at least 100/min, not exceeding 120/min.

With these recommendations being followed, the optimal temporal specifications for the initial five rescue breaths, the circulation check, and the four cardiopulmonary resuscitation (CPR) cycles were estimated and calculated (see Multimedia Appendix 3). The optimal total time was also estimated for the whole sequence, from safety check to emergency call. We scored 2 points for each procedural step if it was performed within ±10% of the optimal estimated calculated time and 1 point if within ±20%. If the participant took a longer or shorter time, no points were scored per step. Two raters measured these times on video recordings. A total of 8 points could be achieved.

Performance: Procedural Quality

Two group-blinded video raters with expertise in PBLS scored the procedural quality of the participants' PBLS skills. The scores were averaged for further analysis. We used a scoring form in trichotomous fashion, with 2 points for correct performance, 1 point for minor deficits, and no points for major deficits (see Multimedia Appendix 4). A maximum of 22 points could be achieved; items were not weighted. Such kinds of scoring systems with comparable checklists are established to assess clinical performances in simulated emergency scenarios [24-27]. In contrast to published rating modalities, we rated the aspects of the algorithm and time measures separately as described above to achieve more objective scoring. In addition, skills performance levels were rated globally: competent, borderline, not competent. Only the performances that were rated "competent" concordantly by both raters were counted and used in the analyses.

Self-Assessments

We developed a self-assessment instrument consisting of seven items on procedural knowledge and seven items on procedural skills (see Multimedia Appendix 5). Two senior pediatricians with expertise in both PBLS and questionnaire design had reviewed these items. Answers were given on 100 mm visual analog scales from 0 (very little confidence) to 100 (highly confident).

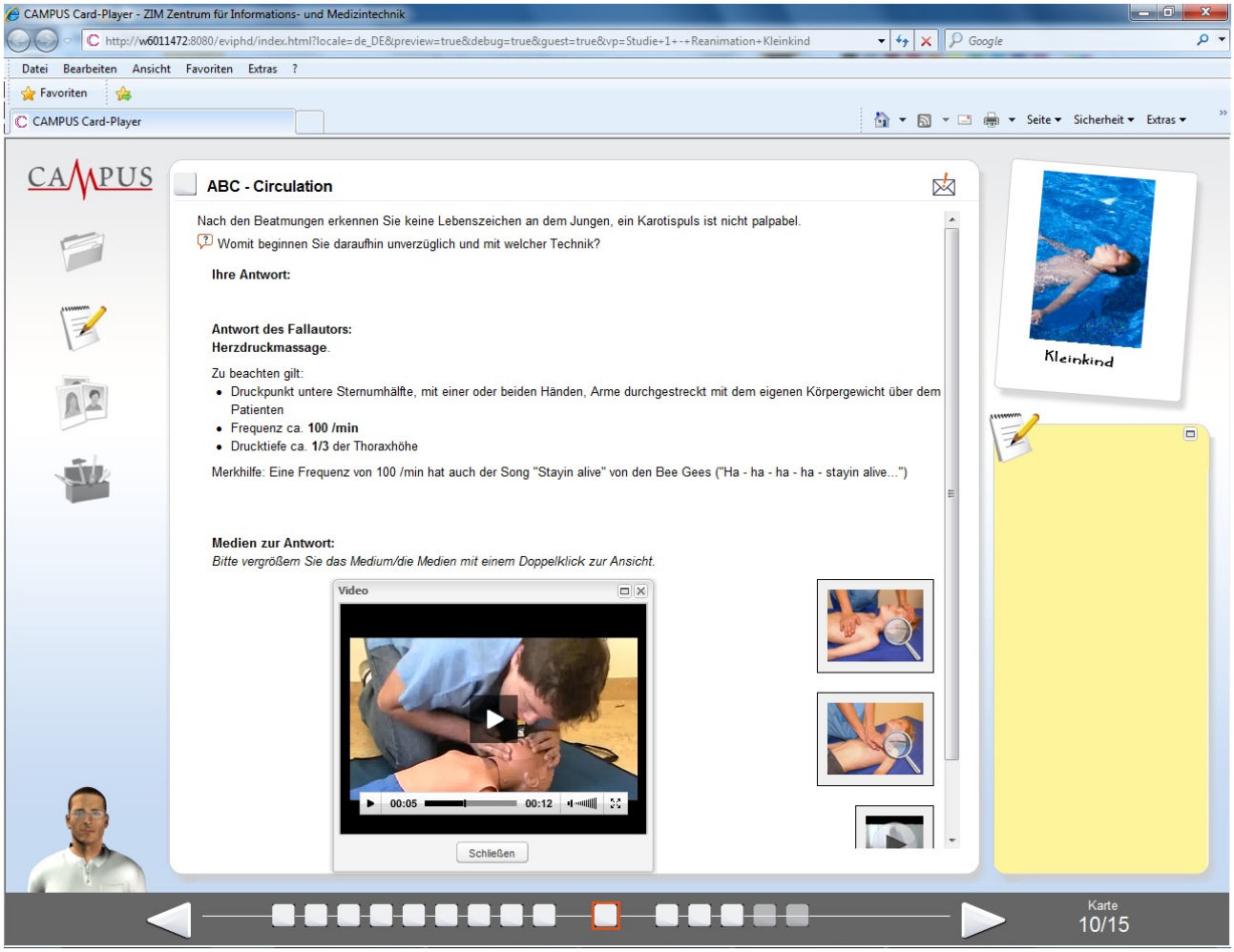
Preparation Material and Pediatric Basic Life Support Training

For individual preparation of the training, we developed and distributed to both groups a paper handout on PBLS. Such handouts are commonly used as preparation for undergraduate skills laboratories [28]. The handout contained all relevant information, explaining the procedural steps of PBLS, including the algorithm, temporal demands, and a flowchart. Additionally, the intervention group was given Web-based access to two VPs dealing with PBLS in infants and toddlers. The VPs were designed with CAMPUS-Software [29] according to published design criteria [11] and enriched by video clips and interactive graphics (see Figure 2). For more detailed characterization of the VP cases used for this study, see Lehmann et al (VP3 and VP4) [17]. Both VPs had to be worked up twice, which was checked electronically but without the ability to identify any participant. The required overall workup time was estimated at 30 to 60 min based on previously measured log data.

Participants were trained in a single-rescuer scenario: from finding an unresponsive child, to the emergency call after 1 min (about four cycles) of CPR according to current guidelines [21]. The hands-on training was divided into two sessions—infant and toddler phases—of 30 min each. The sessions were structured with a commonly used four-step approach [1,28,30-32]. Steps three and four—tutor guided by learner and demonstration by the learner, respectively—were performed once per session by each participant so there was a standardized and comparable amount of individual training time. Two senior tutors provided close feedback on the participants' performance as suggested by Issenberg et al for effective learning during simulations [33]. We used manikins by Laerdal Medical GmbH, Puchheim, Germany ("Baby Anne") and Simulaids Inc, Saugerties, New York ("Kyle").

XSL•FO
RenderX

Figure 2. Screenshot of CAMPUS-Software showing a virtual patient.



Participants and Data Collection

The Ethics Committee of the Medical Faculty Heidelberg granted ethical approval for this study (EK No. S-282/2014). All collected data were pseudonymized. We affirmed with participants by written informed consent that their participation was voluntary, that they could not be identified from the collected data, and that no plausible harm could arise from participation in the study.

We offered participation in this study to a total of about 480 third- and fourth-year medical students at Heidelberg Medical School by group emails and bulletin boards. Invited students had already completed basic life support (BLS) training but had had no PBLS training yet. Announcements were worded as invitations to a special PBLS course and educational study without mentioning e-learning in particular. At an orientation meeting, prospective students enrolled themselves onto a numbered list, unaware of group allocation, which was randomly distributed by numbers.

Rater Selection and Training

We selected and trained two raters to score videotaped performances with the help of best-practice videos of senior faculty. Rater training included reviewing the case content and objectives, and an introduction to the rating schemes. Videotaped examples with different levels of procedural quality were discussed for calibration of the intended use of the schemes. We chose a senior pediatric consultant and a pediatric intensive care nurse practitioner, each of whom was an experienced facilitator for pediatric emergency simulations. Raters were blinded to group classification of all video records.

Data Analysis

Results are presented as the mean ± standard deviation per group and given as the percent of the maximum achievable scores. Data were checked for normal distribution using the Kolmogorov-Smirnov test. If a presumed normal distribution was accepted, statistical differences were evaluated using the unpaired *t* test for between-group comparisons and the paired *t* test for within-group comparisons. Otherwise, we used the Mann-Whitney U test for nonnormal distributions. We assumed that a group difference of 1 SD or more was a relevant effect size. For a group of 30 subjects, we estimated a 65% power to detect this effect, assuming a two-sided significance level of .05. The interrater reliability was estimated using the case 2 intraclass correlation coefficient (ICC2) measured on 100% of the sample size [34]. Global competence-level ratings were compared using Fisher's exact test. As a higher level of qualification in PBLS appeared to be a possible confounder, we confirmed all statistics with exclusion of participants with PBLS qualifications who were identified from the basic data. We used SPSS Statistics version 21 (IBM Corporation, Armonk, NY, USA) for all statistical analyses and an alpha level of .05.

XSL•FO
RenderX

Results

Overview

Scoring results are depicted in Table 1 and Figure 3.

Table 1. Key-feature test, performance, and self-assessment scores.

Scored items	Scores, mean (SD) or n (%)		*P* (CG vs IG)
	Control group (CG)	Intervention group (IG)	
Procedural knowledge: Key-feature test (%), mean (SD)			
t_0[a]	31.0 (12.9)	34.8 (17.1)	.34
t_1[b]	68.8 (16.3)	92.2 (4.7)	*<.001*[c]
P (t_0 vs t_1)	*<.001*	*<.001*	
Performance: Adherence to algorithm (%), mean (SD)			
t_1	72.0 (17.7)	93.4 (7.1)	*<.001*
t_2[d]	95.7 (7.2)	99.8 (1.1)	*.008*
P (t_1 vs t_2)	*<.001*	*<.001*	
Performance: Temporal demands (%), mean (SD)			
t_1	43.3 (23.4)	67.6 (21.4)	*<.001*
t_2	55.8 (27.8)	82.4 (17.8)	*<.001*
P (t_1 vs t_2)	*.03*	*.004*	
Total time of PBLS[e] sequence in seconds, mean (SD)			
t_1	107.7 (35.2)	88.1 (12.6)	*.008*
t_2	95.2 (16.2)	78.1 (10.2)	*<.001*
P (t_1 vs t_2)	.05	*<.001*	
Performance: Procedural quality (%), mean (SD)			
t_1	48.8 (20.2)	68.2 (15.0)	*<.001*
t_2	84.4 (11.2)	89.4 (9.2)	.07
P (t_1 vs t_2)	*<.001*	*<.001*	
Global ratings (rated "competent"), n (%)			
t_1	0/30 (0)	5/27 (19)	*.02*
t_2	17/30 (57)	23/27 (85)	*.02*
Self-assessment (%), mean (SD)			
t_0	29.1 (16.3)	27.2 (18.8)	.68
t_1	59.6 (15.8)	72.3 (11.7)	*.001*
t_2	87.4 (8.6)	88.4 (7.8)	.63

[a]Prepreparation assessment (t_0).

[b]Postpreparation assessment (t_1).

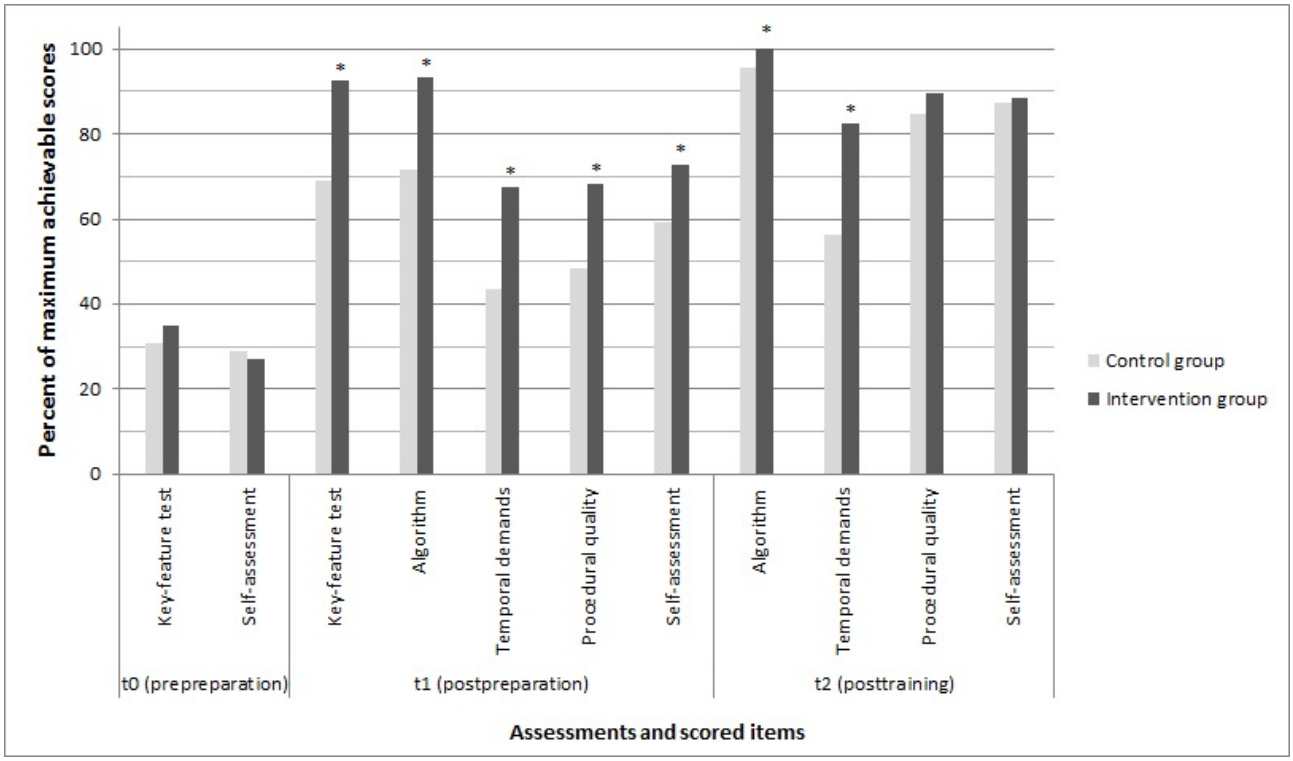
[c]Italicized *P* values represent significant results.

[d]Posttraining assessment (t_2).

[e]Pediatric basic life support (PBLS).

XSL•FO
RenderX

Figure 3. Key-feature test, performance, and self-assessment scores. Scores are given as the percent of the maximum achievable scores (*P<.05).



Basic Data

A total of 57 participants completed the training and all surveys were included in this study—30 (53%) in the control group (CG) and 27 (47%) in the intervention group; approximately 11.9% (57/480) of all eligible students. Out of 60 initial participants, 3 (5%) were excluded due to nonappearance at the training session; all participants of the intervention group processed the VPs completely as requested. Participants' mean age was 24.2 years (SD 2.6) (16/30, 53% female) in the control group and 24.1 years (SD 3.1) (17/27, 63% female) in the intervention group. Of the 57 participants, there were 5 out of 30 (17%) PBLS-qualified participants (paramedics) in the control group and 4 out of 27 (15%) in the intervention group.

Clinical Decision-Making Skills and Procedural Knowledge

There was no significant difference in the key-feature test results between the control group and intervention group at t_0 (31.0%, SD 12.9 vs 34.8%, SD 17.1; P=.34). The intervention group showed a significantly superior increase in procedural knowledge at t_1 compared with the control group (92.2%, SD 4.7 vs 68.8%, SD 16.3; P<.001). There were significant improvements in both groups between t_0 and t_1 (both P<.001).

Performance: Adherence to Algorithm

Regarding adherence to the algorithm, the intervention group was already better than the control group at t_1 (93.4%, SD 7.1 vs 72.0%, SD 17.7; P<.001), which continued at t_2 (99.8%, SD 1.1 vs 95.7%, SD 7.2; P=.008). Significant improvements, however, were found between t_1 and t_2 for both groups (both P<.001).

Performance: Temporal Demands

The intervention group already showed significantly better adherence to temporal specifications than the control group at t_1 (67.6%, SD 21.4 vs 43.3%, SD 23.4; P<.001), which continued at t_2 (82.4%, SD 17.8 vs 55.8%, SD 27.8; P<.001). Both groups showed significant improvements in temporal measures between t_1 and t_2 (P=.03 and P=.004, respectively). Table 1 also shows the measured mean times for the total sequence.

Performance: Procedural Quality

The interrater reliability coefficient was .71 indicating a sufficient level of interrater agreement [35].

The performance quality score of the intervention group was significantly superior to that of the control group at t_1 (68.2%, SD 15.0 vs 48.8%, SD 20.2; P<.001). After practical training, at t_2, they did not differ significantly (89.4%, SD 9.2 vs 84.4%, SD 11.2; P=.07). Both groups showed significantly increased quality scores between t_1 and t_2 (both P<.001).

The global ratings of competence showed significant differences in favor of the intervention group, again already at t_1 and continuing at t_2 (0/30 CG vs 5/27 IG rated "competent", P=.02; 17/30 CG vs 23/27 IG, P=.02, respectively). In all, 85% (23/27) of the intervention group participants performed PBLS "competently" after having practiced with VPs and undergoing PBLS training, compared with only 57% (17/30) of the control group participants.

Self-Assessments

There was no significant difference in the self-assessment means of the two groups at t_0 (29.1%, SD 16.3 CG vs 27.2%, SD 18.8 IG; P=.69). After different preparations, the intervention group

XSL•FO
RenderX

showed a significant increase in its self-assessment compared with that in the control group at t_1 (72.3%, SD 11.7 vs 59.6%, SD 15.8; *P*=.001). At t_2, there was no significant difference between the groups (87.4%, SD 8.6 CG vs 88.4%, SD 7.8 IG; *P*=.62).

Subgroup Analyses

After identifying and excluding PBLS-qualified participants, there were no changes in statistical significances in any of the calculations. The level of significance did not differ by the power of 10.

Discussion

Principal Findings

In this randomized controlled trial, we investigated the impact of an additional preparation with VPs on the improvement of objective and subjective learning outcomes of skill acquisition when combined with standard PBLS training. The control group and intervention group were comparable in terms of their self-assessment and procedural knowledge during the prepreparation assessment. However, after addition of practical training, the intervention group demonstrated significantly better performance in key aspects of PBLS than did the control group, although self-assessment ratings were similar. Also, after practicing with VPs, the intervention group had already demonstrated superior skills, even before the hands-on training in terms of objective skill performance, procedural knowledge, and self-assessment.

Objective Learning Outcomes

After using VPs as an interactive preparation, the intervention group showed significantly improved clinical decision-making skills and procedural knowledge. Also, their PBLS skill performance was superior to that of the control group after the preparation period in regard to objective performance measures, including adherence to the algorithm, temporal demands, and procedural quality. This is in line with existing reports that such electronic learning activities improve both knowledge and skills [18,19]. De Vries et al also showed that a comparable computer simulator improves procedural skills [18]. Although they reported that some of the skill outcomes were suboptimal, the training was not blended with hands-on training as presented here, which led to increased improvements compared with using the computer simulator alone. Furthermore, as reported by Ventre et al, such approaches might fill a gap in continuing medical education [20]. Procedural skill performance was rated as objectively as possible to discriminate procedural learning effects. In contrast, typically used checklists often subsume adherence to the algorithm, temporal aspects, and performance quality—for example, “CPR continued—2 points for initiated immediately after pulse check and rhythm identification (<30 s) *and* good CPR technique *and* checks pulse with CPR” (taken from The Clinical Performance Tool [26]). At t_2, when both groups had had equal practical training, the procedural steps of PBLS were still performed qualitatively more competently by the intervention group in some aspects. Such differences will probably not be found when using global rating scales, but may be when using automated skill reporting devices as used by Kononowicz et al [19], or when using discriminating rating schemes as presented here when such devices are not available.

We assume that VPs facilitated application of acquired clinical decision-making skills and procedural knowledge. Interactivity and feedback in VPs, which included interactive graphics and video clips, might have enhanced the learning process beyond the use of media, as in other approaches. It is well known that educational feedback, such as that given in the VPs, is the most important feature of simulation-based education [33]. Interacting with clinical case scenarios might also provide an emotionally activating stimulus to get trainees involved as it supports the acquisition and retention of skills [36]. For complex procedures, current learning theories support a reasonable simple-to-complex learning process that facilitates learning [37,38]. VPs may bridge this gap between knowledge and practice.

The presented results support the subjective perceptions of students and tutors [17] that such a blended learning approach is effective and efficient for procedural learning. In this study, self-directed learning with paper handouts seems to have had little effect on facilitating the acquisition of practical skills, although it did have an effect on improving procedural knowledge. In contrast, the blended approach that included interactive VPs for preparation led to improved learning of both procedural knowledge and procedural skills. Implications for CPR and other emergency training might be a more efficient and effective use of resource-intensive training time.

Subjective Learning Outcomes

In their self-assessments, the participants of the intervention group judged themselves superior to those in the control group after the preparation period. Objective findings in their scored performances support these ratings. After their practical training, however, the self-assessments of the two groups were similar. In contrast, the intervention group still had superior objective scores regarding skill performance. Self-assessments are not necessarily correlated with performance; for example, postgraduate practitioners have limited ability to self-assess accurately, as shown by Davis et al [39].

Study Strengths

The assessments of clinical decision-making skills and procedural knowledge, practical performance, and self-assessment combine relevant and detailed objective and subjective measures for elucidating the learning effects of this approach. This is one of the first studies that provides objective data that support how effectively VPs can foster the acquisition of PBLS skills.

Study Limitations

Participants’ VP case completions were monitored to validate their workup, but the validation was not done in a controlled environment that allowed evaluation of participants’ efforts. Accordingly, efforts on the workup of handouts were also not assessed. Workup of VPs might have led to more motivation for learning even though the study was not announced as an e-learning study attracting mainly tech-savvy students. Because both groups had significantly increased procedural knowledge after the preparation period, we assumed that motivation for

XSL•FO
RenderX

preparation might not have been very different. Result details, for example, of the temporal scoring, suggest that VPs address skills not affected by paper-based learning materials. However, group differences might also have been influenced by different durations of their efforts to learn in addition to different modalities. Additionally, both groups were not limited in their access to other learning resources than those provided. Also, most of the instruments used in this study have not been validated formally, although all were developed based on current literature and were pilot-tested. Finally, the sample size is rather limited, thereby providing limited power to investigate differences between groups.

Conclusions

The blended learning approach described herein leads to improved outcomes of practical skill acquisition compared with a standard approach. Even before having practical training, preparation with VPs leads to improved practical performances as well as better clinical decision-making skills and procedural knowledge. Further studies are necessary to understand the specific benefit of using VPs regarding clinical skill acquisition and its sustainability.

Acknowledgments

This study was funded by general departmental funds without involving third party funds.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Key-feature test.

[PDF File (Adobe PDF File), 310KB - jmir_v17i7e162_app1.pdf]

Multimedia Appendix 2

Adherence to algorithm scoring form.

[PDF File (Adobe PDF File), 85KB - jmir_v17i7e162_app2.pdf]

Multimedia Appendix 3

Temporal measures scoring form.

[PDF File (Adobe PDF File), 163KB - jmir_v17i7e162_app3.pdf]

Multimedia Appendix 4

Performance quality scoring form.

[PDF File (Adobe PDF File), 151KB - jmir_v17i7e162_app4.pdf]

Multimedia Appendix 5

Self-assessment instrument.

[PDF File (Adobe PDF File), 143KB - jmir_v17i7e162_app5.pdf]

References

1. Bullock I. Skill acquisition in resuscitation. Resuscitation 2000 Jul;45(2):139-143. [Medline: 10950322]
2. Adler MD, Vozenilek JA, Trainor JL, Eppich WJ, Wang EE, Beaumont JL, et al. Development and evaluation of a simulation-based pediatric emergency medicine curriculum. Acad Med 2009 Jul;84(7):935-941. [doi: 10.1097/ACM.0b013e3181a813ca] [Medline: 19550192]
3. Cheng A, Goldman RD, Aish MA, Kissoon N. A simulation-based acute care curriculum for pediatric emergency medicine fellowship training programs. Pediatr Emerg Care 2010 Jul;26(7):475-480. [doi: 10.1097/PEC.0b013e3181e5841b] [Medline: 20577139]
4. Cheng A, Hunt EA, Donoghue A, Nelson K, Leflore J, Anderson J, EXPRESS Pediatric Simulation Research Investigators. EXPRESS--Examining Pediatric Resuscitation Education Using Simulation and Scripting. The birth of an international pediatric simulation research collaborative--from concept to reality. Simul Healthc 2011 Feb;6(1):34-41. [doi: 10.1097/SIH.0b013e3181f6a887] [Medline: 21330848]

XSL•FO
RenderX

5. Lynagh M, Burton R, Sanson-Fisher R. A systematic review of medical skills laboratory training: where to from here? Med Educ 2007 Sep;41(9):879-887. [doi: 10.1111/j.1365-2923.2007.02821.x] [Medline: 17696985]
6. Carrero E, Gomar C, Penzo W, Fábregas N, Valero R, Sánchez-Etayo G. Teaching basic life support algorithms by either multimedia presentations or case based discussion equally improves the level of cognitive skills of undergraduate medical students. Med Teach 2009 May;31(5):e189-e195. [doi: 10.1080/01421590802512896] [Medline: 19241215]
7. Braslow A, Brennan RT, Newman MM, Bircher NG, Batcheller AM, Kaye W. CPR training without an instructor: development and evaluation of a video self-instructional system for effective performance of cardiopulmonary resuscitation. Resuscitation 1997 Jun;34(3):207-220. [Medline: 9178381]
8. Sopka S, Biermann H, Rossaint R, Knott S, Skorning M, Brokmann JC, et al. Evaluation of a newly developed media-supported 4-step approach for basic life support training. Scand J Trauma Resusc Emerg Med 2012;20:37 [FREE Full text] [doi: 10.1186/1757-7241-20-37] [Medline: 22647148]
9. Todd KH, Braslow A, Brennan RT, Lowery DW, Cox RJ, Lipscomb LE, et al. Randomized, controlled trial of video self-instruction versus traditional CPR training. Ann Emerg Med 1998 Mar;31(3):364-369. [Medline: 9506495]
10. Cook DA, Triola MM. Virtual patients: a critical literature review and proposed next steps. Med Educ 2009 Apr;43(4):303-311. [doi: 10.1111/j.1365-2923.2008.03286.x] [Medline: 19335571]
11. Huwendiek S, Reichert F, Bosse H, de Leng BA, van der Vleuten CP, Haag M, et al. Design principles for virtual patients: a focus group study among students. Med Educ 2009 Jun;43(6):580-588. [doi: 10.1111/j.1365-2923.2009.03369.x] [Medline: 19493183]
12. Iserbyt P, Charlier N, Mols L. Learning basic life support (BLS) with tablet PCs in reciprocal learning at school: are videos superior to pictures? A randomized controlled trial. Resuscitation 2014 Jun;85(6):809-813. [doi: 10.1016/j.resuscitation.2014.01.018] [Medline: 24503451]
13. Burnette K, Ramundo M, Stevenson M, Beeson MS. Evaluation of a web-based asynchronous pediatric emergency medicine learning tool for residents and medical students. Acad Emerg Med 2009 Dec;16 Suppl 2:S46-S50. [doi: 10.1111/j.1553-2712.2009.00598.x] [Medline: 20053211]
14. Roe D, Carley S, Sherratt C. Potential and limitations of e-learning in emergency medicine. Emerg Med J 2010 Feb;27(2):100-104. [doi: 10.1136/emj.2008.064915] [Medline: 20156859]
15. Shah IM, Walters MR, McKillop JH. Acute medicine teaching in an undergraduate medical curriculum: a blended learning approach. Emerg Med J 2008 Jun;25(6):354-357. [doi: 10.1136/emj.2007.053082] [Medline: 18499818]
16. Spedding R, Jenner R, Potier K, Mackway-Jones K, Carley S. Blended learning in paediatric emergency medicine: preliminary analysis of a virtual learning environment. Eur J Emerg Med 2013 Apr;20(2):98-102. [doi: 10.1097/MEJ.0b013e3283514cdf] [Medline: 22415016]
17. Lehmann R, Bosse HM, Simon A, Nikendei C, Huwendiek S. An innovative blended learning approach using virtual patients as preparation for skills laboratory training: perceptions of students and tutors. BMC Med Educ 2013;13:23 [FREE Full text] [doi: 10.1186/1472-6920-13-23] [Medline: 23402663]
18. de Vries W, Handley AJ. A web-based micro-simulation program for self-learning BLS skills and the use of an AED. Can laypeople train themselves without a manikin? Resuscitation 2007 Dec;75(3):491-498. [doi: 10.1016/j.resuscitation.2007.05.014] [Medline: 17629390]
19. Kononowicz AA, Krawczyk P, Cebula G, Dembkowska M, Drab E, Frączek B, et al. Effects of introducing a voluntary virtual patient module to a basic life support with an automated external defibrillator course: a randomised trial. BMC Med Educ 2012;12:41 [FREE Full text] [doi: 10.1186/1472-6920-12-41] [Medline: 22709278]
20. Ventre KM, Collingridge DS, DeCarlo D. End-user evaluations of a personal computer-based pediatric advanced life support simulator. Simul Healthc 2011 Jun;6(3):134-142. [doi: 10.1097/SIH.0b013e318207241e] [Medline: 21646982]
21. Biarent D, Bingham R, Eich C, López-Herce J, Maconochie I, Rodríguez-Núñez A, et al. European Resuscitation Council Guidelines for Resuscitation 2010 Section 6. Paediatric life support. Resuscitation 2010 Oct;81(10):1364-1388. [doi: 10.1016/j.resuscitation.2010.08.012] [Medline: 20956047]
22. Kopp V, Möltner A, Fischer M. Key feature problems for the assessment of procedural knowledge: a practical guide. GMS Z Med Ausbild 2006;23(3):Doc50 [FREE Full text]
23. Page G, Bordage G, Allen T. Developing key-feature problems and examinations to assess clinical decision-making skills. Acad Med 1995 Mar;70(3):194-201. [Medline: 7873006]
24. Adler MD, Vozenilek JA, Trainor JL, Eppich WJ, Wang EE, Beaumont JL, et al. Comparison of checklist and anchored global rating instruments for performance rating of simulated pediatric emergencies. Simul Healthc 2011 Feb;6(1):18-24. [doi: 10.1097/SIH.0b013e318201aa90] [Medline: 21330846]
25. Donoghue A, Nishisaki A, Sutton R, Hales R, Boulet J. Reliability and validity of a scoring instrument for clinical performance during Pediatric Advanced Life Support simulation scenarios. Resuscitation 2010 Mar;81(3):331-336. [doi: 10.1016/j.resuscitation.2009.11.011] [Medline: 20047787]
26. Donoghue A, Ventre K, Boulet J, Brett-Fleegler M, Nishisaki A, Overly F, EXPRESS Pediatric Simulation Research Investigators. Design, implementation, and psychometric analysis of a scoring instrument for simulated pediatric resuscitation: a report from the EXPRESS pediatric investigators. Simul Healthc 2011 Apr;6(2):71-77. [doi: 10.1097/SIH.0b013e31820c44da] [Medline: 21358564]

XSL•FO
RenderX

27. Lockyer J, Singhal N, Fidler H, Weiner G, Aziz K, Curran V. The development and testing of a performance checklist to assess neonatal resuscitation megacode skill. Pediatrics 2006 Dec;118(6):e1739-e1744. [doi: 10.1542/peds.2006-0537] [Medline: 17074839]
28. Krautter M, Weyrich P, Schultz J, Buss SJ, Maatouk I, Jünger J, et al. Effects of Peyton's four-step approach on objective performance measures in technical skills training: a controlled trial. Teach Learn Med 2011;23(3):244-250. [doi: 10.1080/10401334.2011.586917] [Medline: 21745059]
29. Centre for Virtual Patients. CAMPUS-Software URL: http://www.medizinische-fakultaet-hd.uni-heidelberg.de/CAMPUS-Software.109992.0.html?&FS=0&L=en [accessed 2015-06-25] [WebCite Cache ID 6ZYS3m5w6]
30. Cummins RO. Advanced Cardiovascular Life Support, Instructor's Manual (American Heart Association). Dallas, TX: American Heart Association; 2001.
31. Nolan J, Baskett P, Gabbott D, Gwinnutt C, Latorre F, Lockey A. In: Nolan J, editor. Advanced Life Support Course: Provider Manual. 4th edition. London, UK: Resuscitation Council; 2001.
32. Peyton JWR, editor. Teaching & Learning in Medical Practice. Rickmansworth, UK: Manticore Publishers; 1998.
33. Issenberg SB, McGaghie WC, Petrusa ER, Lee GD, Scalese RJ. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. Med Teach 2005 Jan;27(1):10-28. [doi: 10.1080/01421590500046924] [Medline: 16147767]
34. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979 Mar;86(2):420-428. [Medline: 18839484]
35. Fleiss JL. The Design and Analysis of Clinical Experiments. New York, NY: Wiley-Interscience; 1986.
36. Beckers SK, Biermann H, Sopka S, Skorning M, Brokmann JC, Heussen N, et al. Influence of pre-course assessment using an emotionally activating stimulus with feedback: a pilot study in teaching Basic Life Support. Resuscitation 2012 Feb;83(2):219-226. [doi: 10.1016/j.resuscitation.2011.08.024] [Medline: 21924220]
37. Sweller J, van Merrienboer JJG, Paas FGWC. Cognitive architecture and instructional design. Educ Psychol Rev 1998 Sep;10(3):251-296 Print ISSN 1040-726X. [doi: 10.1023/A:1022193728205]
38. van Merriënboer JJ, Sweller J. Cognitive load theory in health professional education: design principles and strategies. Med Educ 2010 Jan;44(1):85-93. [doi: 10.1111/j.1365-2923.2009.03498.x] [Medline: 20078759]
39. Davis DA, Mazmanian PE, Fordis M, Van Harrison R, Thorpe KE, Perrier L. Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. JAMA 2006 Sep 6;296(9):1094-1102. [doi: 10.1001/jama.296.9.1094] [Medline: 16954489]

Abbreviations

BLS: basic life support
CG: control group
CPR: cardiopulmonary resuscitation
ICC2: case 2 intraclass correlation coefficient
IG: intervention group
PBLS: pediatric basic life support
t0: prepreparation assessment
t1: postpreparation assessment
t2: posttraining assessment
VP: virtual patient

Edited by G Eysenbach; submitted 13.12.14; peer-reviewed by A Kononowicz, P Iserbyt; comments to author 03.03.15; revised version received 14.04.15; accepted 24.05.15; published 02.07.15.

Please cite as:
Lehmann R, Thiessen C, Frick B, Bosse HM, Nikendei C, Hoffmann GF, Tönshoff B, Huwendiek S
Improving Pediatric Basic Life Support Performance Through Blended Learning With Web-Based Virtual Patients: Randomized Controlled Trial
J Med Internet Res 2015;17(7):e162
URL: http://www.jmir.org/2015/7/e162/
doi:10.2196/jmir.4141
PMID:26139388

©Ronny Lehmann, Christiane Thiessen, Barbara Frick, Hans Martin Bosse, Christoph Nikendei, Georg Friedrich Hoffmann, Burkhard Tönshoff, Sören Huwendiek. Originally published in the Journal of Medical Internet Research (http://www.jmir.org),

XSL•FO
RenderX

02.07.2015. 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

Scientific Versus Experiential Evidence: Discourse Analysis of the Chronic Cerebrospinal Venous Insufficiency Debate in a Multiple Sclerosis Forum

Janka Koschack[1], PhD; Lara Weibezahl[1]; Tim Friede[2], PhD; Wolfgang Himmel[1], PhD; Philip Makedonski[3], MSc; Jens Grabowski[3], PhD

[1]Department of General Practice, University Medical Center Göttingen, Göttingen, Germany

[2]Department of Medical Statistics, University Medical Center Göttingen, Göttingen, Germany

[3]Institute of Computer Science, Georg-August- University Göttingen, Göttingen, Germany

Corresponding Author:
Janka Koschack, PhD
Department of General Practice
University Medical Center Göttingen
Humboldtallee 38
Göttingen, 37073
Germany
Phone: 49 5513914225
Fax: 49 551399530
Email: jkoscha@gwdg.de

Abstract

Background: The vascular hypothesis of multiple sclerosis (MS), called chronic cerebrospinal venous insufficiency (CCSVI), and its treatment (known as liberation therapy) was immediately rejected by experts but enthusiastically gripped by patients who shared their experiences with other patients worldwide by use of social media, such as patient online forums. Contradictions between scientific information and lay experiences may be a source of distress for MS patients, but we do not know how patients perceive and deal with these contradictions.

Objective: We aimed to understand whether scientific and experiential knowledge were experienced as contradictory in MS patient online forums and, if so, how these contradictions were resolved and how patients tried to reconcile the CCSVI debate with their own illness history and experience.

Methods: By using critical discourse analysis, we studied CCSVI-related posts in the patient online forum of the German MS Society in a chronological order from the first post mentioning CCSVI to the time point when saturation was reached. For that time period, a total of 117 CCSVI-related threads containing 1907 posts were identified. We analyzed the interaction and communication practices of and between individuals, looked for the relation between concrete subtopics to identify more abstract discourse strands, and tried to reveal discourse positions explaining how users took part in the CCSVI discussion.

Results: There was an emotionally charged debate about CCSVI which could be generalized to 2 discourse strands: (1) the "downfall of the professional knowledge providers" and (2) the "rise of the nonprofessional treasure trove of experience." The discourse strands indicated that the discussion moved away from the question whether scientific or experiential knowledge had more evidentiary value. Rather, the question whom to trust (ie, scientists, fellow sufferers, or no one at all) was of fundamental significance. Four discourse positions could be identified by arranging them into the dimensions "trust in evidence-based knowledge," "trust in experience-based knowledge," and "subjectivity" (ie, the emotional character of contributions manifested by the use of popular rhetoric that seemed to mask a deep personal involvement).

Conclusions: By critical discourse analysis of the CCSVI discussion in a patient online forum, we reconstruct a lay discourse about the evidentiary value of knowledge. We detected evidence criteria in this lay discourse that are different from those in the expert discourse. But we should be cautious to interpret this dissociation as a sign of an intellectual incapability to understand scientific evidence or a naïve trust in experiential knowledge. Instead, it might be an indication of cognitive dissonance reduction to protect oneself against contradictory information.

(J Med Internet Res 2015;17(7):e159) doi:10.2196/jmir.4103

XSL•FO
RenderX

KEYWORDS

multiple sclerosis; venous insufficiency; Internet; social media; cognitive dissonance; qualitative research

Introduction

Patients increasingly search the Internet for information on medical conditions, including clinical news and treatment options. Although this information is typically provided by medical experts or commercial sources, there has also been an increase in peer-to-peer health care [1]. Not only evidence from scientific sources, such as the latest results of clinical trials, diffuses into the lay community in this way. Patients also share the experiences they have with doctors, treatments, and the everyday living with a disease with other patients in online forums [2]. These shared patient experiences have formed a new database of experiential knowledge that is not only a source of information for patients and their relatives, but also has increasing relevance for scientific research [3].

The patients' use of the Internet as a source of both scientific and experiential knowledge is a cause of serious concern when these different forms of knowledge do not peacefully coexist, but are contradictory. This was recently observed in the debate about the endovascular treatment of multiple sclerosis (MS), which is based on a new etiologic hypothesis called chronic cerebrospinal venous insufficiency (CCSVI) [4]. Although the scientific community was opposed to or ignored the CCSVI hypothesis, it was heatedly debated in online patient communities, particularly the resulting treatment (ie, modified venous angioplasty or stenting of jugular and azygous veins) [5]. Although scientists and clinicians strongly advised against this procedure before it was rigorously scientifically examined for efficacy and safety [6], patients all over the world found a way to obtain access to this procedure, often commercially referred to as the "liberation treatment" [7]. Patients' experiences with the liberation treatment were soon published on the popular video-sharing website YouTube. Mazanderani and colleagues [8] showed in a content analysis of these YouTube videos that patients used the videos to prove the effectiveness of the treatment, for instance, by showing improved symptoms after partaking in the treatment.

The debate about the CCSVI hypothesis and the associated intervention took place in countless patient forums all over the world. However, particularly in Canada, the demand of the patient community and its advocates for further research was so strong that CCSVI and the liberation treatment became a research topic, not defined by the scientific community itself, but by the patient community [9]. The CCSVI hype has abated somewhat since the most recent studies found neither evidence for a high prevalence of CCSVI nor for a causal relationship to MS, and a wave of complications following venous stenting and angioplasty was reported instead [10].

The CCSVI story is now seen as a "waste of valuable time, money, and intellectual energy" [11], at least by large parts of the scientific community. We know that the Internet and Web 2.0 played an important role in mobilizing thousands of participants in the CCSVI debate [12]. What we do not know is whether patients felt a conflict between their own understanding of evidentiary value and the agenda of the scientific community during the CCSVI debate and, if so, how they reconciled conflicting opinions, interests, and objectives.

By analyzing the CCSVI discussion in the patient online forum of the German MS Society (Deutsche Multiple Sklerose Gesellschaft; DMSG), we aimed to reconstruct the underlying discourse that forms this discussion. One approach to uncover discourses was introduced by the Duisburg School of Critical Discourse Analysis [13]. It is based on Michel Foucault's discourse theory that deals with questions such as what knowledge is, how the valid knowledge evolves, which function it has for the constitution of subjects, and the shaping of society. Knowledge means all kinds of meanings used by real persons to interpret and shape the surrounding reality.

Methods

Database General Description

The database for this analysis consisted of contributions posted in the online forum of the DMSG. The DMSG presents itself on its website as a nonprofit stakeholder of MS patients and their relatives. It is a registered charity with more than 900 community contact groups. Among other features, the DMSG provides 2 different types of freely accessible forums on its website: an expert forum with time-limited chats between experts and users about different issues (eg, cognitive deficits or pregnancy) and a second forum that is unstructured, not moderated, and open for anonymous registration. It is targeted at laypeople, mostly patients with MS. The forum consists of threads. A thread is a . These postings can contain hyperlinks and can cite any number of previous postings. Every user can open a new thread or contribute to an existing one.

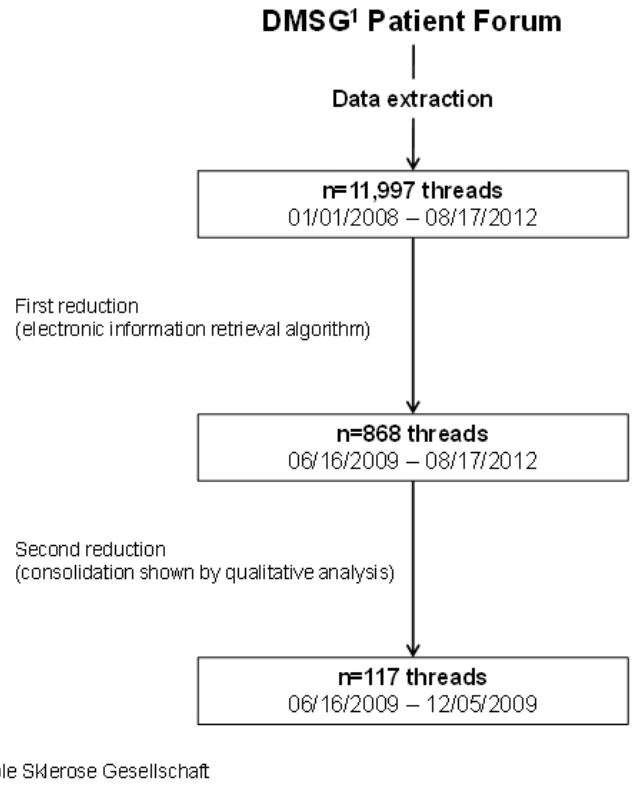
Data Reduction

Initially, all contributions between January 1, 2008 (the starting point of the forum) and August 17, 2012 (the date of the extraction) were extracted. This initial database consisted of 139,912 postings and was reduced first to postings contributing to the CCSVI discussion. The information retrieval algorithm for identifying individual postings is described in detail for the quantitative analysis [14]. A total of 868 CCSVI-related threads containing more than 53,000 postings were identified. The threads varied by numbers of postings; a few contained only one posting and the longest had more than 2000 postings. The first CCSVI-related posting determined the beginning of our data analysis. To define the end of the chronological analysis, we followed the concept of saturation, a guiding principle in qualitative research. Our sample had to be large enough to assure that most or all the perceptions that might be important were uncovered, but at the same time if the sample was too large, the data would become redundant. Thus, we looked for a consolidation on levels such as contributors, statements, and argumentations. Six months after the first CCSVI-related posting, we could not identify further discourse fragments that provided new information or put the data already gathered into

XSL•FO
RenderX

perspective. The flowchart in Figure 1 shows the steps of the data extraction and reduction.

Figure 1. Flowchart of the data extraction and reduction process.



Analysis

Many critical discourse analyses deal with text that is produced under certain formal criteria and related to a specific topic (ie, newspaper articles about the increasing rate of childhood obesity). In this example, the newspaper articles are the textual elements to a particular topic and are called "discourse fragments." These discourse fragments form the "discourse strand" (in this example, the debate about childhood obesity). In our analysis, we defined "CCSVI" as the discourse strand formed by the associated discourse fragments, (ie, the threads related to this topic). Jäger [13] suggests a sequential procedure with 2 main steps: (1) structural analysis of the discourse strand and (2) detailed analysis of typical discourse fragments, including context and surface of the thread, content of the thread, ideological statements in the associated postings, and use of collective symbolism in the associated postings.

Structural Analysis of the Discourse Strand

We read and reread the CCSVI-related threads of the patient forum in a chronological order. We identified significant users (ie, users who posted very often) or whose postings started lively discussions. These users also gave us a preliminary idea of discourse positions that would need to be highlighted and defined in the next step of the analysis. We gathered themes in the discourse that were discussed repeatedly and tried to determine their meaning and impact on the discourse. Three characteristics seemed especially promising for identifying discourse fragments (ie, threads) as typical for the discourse strand: (1) when they started a lively discussion in the forum, meaning that many different users with different opinions reacted to that posting; (2) when they triggered certain (announced or reported) actions in real-life, such as talking with one's own doctor about CCSVI; and (3) when they activated a certain style or nature of the discussion (eg, polemic argumentation).

Detailed Analysis of Typical Discourse Fragments

Context and Surface of the Thread

We described the thread, including the number of contributing users, whether it was from a well-known user (ie, from a user who often posted and elicited many responses from other users), the number of postings, the course of the thread (eg, whether it meandered or was concise), and whether it mentioned subtopics that were discussed previously. We then described the context of the thread, meaning information about whether the thread was triggered by a real-life event. We also reported any characteristics of the thread that might be important for interpreting it.

Content of the Thread

We analyzed the thread chronologically to reveal its logical structure. Subtopics were scrutinized for their meaning (ie, which emotions and connotations were addressed) and how a subtopic was related with other subtopics.

Ideological Statements in the Associated Postings

We analyzed the arguments used to justify attitudes or emotions. Arguments were deconstructed into their superficial content (eg, "the subgroup of primary-progressive MS patients is

XSL•FO
RenderX

neglected") and we flagged the strategy that was used to prove the content (eg, by quoting an expert who stated it). We then determined certain underlying dimensions or features that help reconstruct the different discourse positions from which the forum users evaluated the CCSVI debate.

Use of Collective Symbolism in the Associated Postings

We determined cultural stereotypes, such as figures of speech and allegories, that seemed to be a common ground for the users in the forum, signified something specific, and were therefore able to popularize knowledge.

We followed Jäger's suggestion of a cyclic and iterative working to reveal connections between different aspects of the analysis, to develop interpretations, pinpoint arguments for or against these interpretations, and to reconstruct certain discourse positions from which subjects participate in and evaluate the discourse.

Critical discourse analysis as an interpretative method does not need to present quotations as proof or examples as is commonly done in other qualitative approaches, such as in a content analysis. We only present quotations in one case to illustrate the use of collective symbolism. To demonstrate our methodical approach, Textbox 1 provides an example of the detailed analysis of a typical discourse fragment [15]. The thread was translated from German into English and shortened for demonstration purposes. Main contributors to the qualitative analysis were 3 of the authors (LW, WH, JK). Analysis was done in a team approach, so differences were resolved by in-depth discussion and consensus was sought during the process of analysis. Important intermediate results were discussed with the other 3 authors.

XSL•FO
RenderX

Textbox 1. Detailed analysis of a thread [16] as a typical discourse fragment (thread translated and shortened to the first 2 postings).

Post 1: "Dr XXX's answer to questions concerning venous MS makes the distance abundantly clear that certified doctors have to existential interests of MS patients! The status quo is just right and comfortable. Please do not disturb! I don't know whether the story about the veins is correct, but the lack of willingness to actually apply oneself to this subject clearly shows that there's nothing to be expected from the certified angle, not even disproving seems to be necessary. I sincerely thank you, Dr XXX, for this honest confession (a bitter truth for us). I've never heard it admitted this openly." (User 1, 11:40 am September 18, 2009)

Post 2: "These gods in white don't know a thing, they just read the press releases of the pharma mafia about just how potent these dubious BT are." (User 2, 00:19 pm September 18, 2009)

Thread title: "The overstrained/threadbare expert" [German: "Der überstrapazierte Experte"]

Time period: First posting 11:40 am on September 18, 2009; last posting 2:25 pm on September 19, 2009

Participants: 18 users; several well-known users (regular contributors) of the forum (eg, User 1, User 2 [usernames replaced with placeholders])

Thread characteristics: 28 postings; a short, but very dense thread dealing with the following subtopics: "who is the expert," "Big Pharma," "what constitutes trustworthiness," and "DMSG no patients advocate"

Context and content: Thread is about an event from outside the patient forum, but within the DMSG website: Three days before (8:02 pm September 15, 2009), the editorial office deleted some postings addressed to an expert together with the expert's answers. The office explained this removal as follows: The postings were not dealing with the announced topic of the expert forum (ie, "Different courses—different therapies"). Instead, the users consulted the expert about CCSVI.

Description (of the first 2 postings): User 1 (well-known; patient with primary-progressive MS [PPMS]; already suffering from some disabilities) starts with a quotation of Dr. XXX [not printed here]. He interprets Dr. XXX's statement as exemplary for the dissociation between the needs of the patient community and the attitudes of the expert community. He uses the attribute "certified" for the medical profession—although a doctor is qualified per se, there is no additional certification needed. User 1 stated his frustration about the expert's lack of interest to discuss CCSVI and he interpreted this lack of interest less as an opinion about CCSVI, but more as the typical arrogant stance experts have toward patients. Thirty minutes later, User 2 responded to User 1′s posting (User 2 is also well-known; MS patient with PPMS; already suffering from some disabilities). User 2 uses the [German] phrase "gods in white," which expresses the widely shared opinion that doctors are almost almighty, but they do not share the afflictions of common people at all. "White" in this phrase refers to the color of the commonly worn doctor's coat. His negative judgment is fueled by the postpositive statement that doctors are naive because they uncritically believe the pharmaceutical industry's advertising of "BT" (BT is the abbreviation of beta-interferon, the active ingredient in the currently most commonly prescribed medication for relapse-remitting MS in Germany).

Comments (about the first 2 postings): The thread title uses a German phrase that can be interpreted in 2 ways: as "overstrained" (meaning that the expert is unable to answer adequately for various reasons) vs "threadbare" (ie, the expert is not an expert at all, his declared status as an expert is a farce).

The thread title already shows the ironically contested expert status of someone who is in fact an expert; thus, expertise as a sign of quality for health care is doubted (subtopic "who is the expert").

User 1 must assume that he is perceived as trustworthy by the other users because the posts are already deleted, it is not possible to determine whether this is true—trustworthiness as a matter of personal involvement and being recognizable as an individual (subtopic "what constitutes trustworthiness").

By doubting Dr. XXX's expert status, the expert status of the DMSG is also disputed because Dr. XXX is a member of the DMSG advisory board (subtopic "DMSG no patients advocate")

User 2 shows by using the abbreviation "BT," which only insiders know about, that he considers himself some kind of expert with special knowledge (subtopic "who is the expert").

Doctors are not seen predominantly as selfish betrayers, but as being caught by their own arrogance not to see that they are as framed as the patients by the pharmaceutical industry (subtopics "Big Pharma" and "who is the expert").

Collective symbolism (in the first 2 postings): "God in white": a stylistic device to ironically qualify doctors as arrogantly believing being capable of everything (ie, like a god).

Ethical Consideration

There are no rules for the ethical challenge that is inherent to using health discussion board postings as research data [16]. Because we were unable to obtain informed consent from the forum visitors to use the data they produced by posting to the forum, we officially informed the executive board of the DMSG about the study and they gave us their consent. Additionally, we asked the Ethics Committee of the University Medical Center Göttingen for approval. The committee decided that an approval was not necessary (11/5/13An). In the rare cases of using quotations, we used pseudonyms instead of real usernames or replaced the real names with XXX, respectively.

Results

The first posting related to CCSVI was published on June 16, 2009, and we completed our analysis with data from the end of 2009. At that time, the CCSVI discourse consisted of 117 threads (1907 postings) (see Figure 1). We identified 2 main discourse strands. Some collective symbols attracted our attention during the detailed analysis. Finally, we disclosed certain dimensions of argumentation and reconciliation that apparently were useful for interpreting discourse positions within the CCSVI discourse.

XSL•FO
RenderX

Discourse Strands

We identified certain subtopics in the threads, such as "who is the expert," "Big Pharma," and "what constitutes trustworthiness" (see Textbox 1). By determining the relationship between these subtopics, we reconstructed 2 discourse strands: (1) the "downfall of the professional knowledge providers" and (2) the "rise of the nonprofessional treasure trove of experience."

The Downfall of the Professional Knowledge Providers

The first posting mentioning CCSVI was answered 5 hours later with a link to the original CCSVI study. Like scientists, some of the users began to discuss the CCSVI hypothesis and Zamboni's study [4] against the background of evidence-based medicine, using terms such as "placebo," "number of cases," and "blinded." However, this parascientific discourse was not continued; the lack of any further evidence-based information may be the reason. Although users clearly understood that it was necessary to have more evidence, some of them argued that the progressive course of their disease did not to allow them to wait until the scientific community produced better knowledge.

Because actual evidence-based knowledge was lacking at that time, the archive of scientific knowledge was examined and an older publication from 1986 was introduced [17]. The long time between the first idea of MS of a vascular disease in 1986 and the new article more than 30 years later caused some users to ask why this theory had not been proven in the past—their answer was that opportunistic interests had blocked and hindered further scientific investigation. Reasonable doubts regarding the scientific validity of the hypothesis were not discussed.

This feeling of distrust of the scientific community or, more precisely, of the well-established neurological scientific community tainted the ongoing discussion and the following events were mainly interpreted as confirmation of this feeling. In August 2009, 2 months after the first CCSVI-related posting in the forum, a conference was held by Zamboni himself in Bologna, Italy, raising the users' hope that CCSVI now would become an important topic in the scientific community. This conference was largely ignored by the scientific community, which was interpreted as more proof that mainstream research in MS was not a stakeholder of patient interests. The forum users did not discuss that the conference lacked the common attributes of a scientific convention: only researchers who collaborated with Zamboni took part (ie, no scientists with a negative opinion about CCSVI). Users instead interpreted the situation as evidence for the ignorant mainstream knowledge providers.

The Rise of the Nonprofessional Treasure Trove of Experience

From the beginning of the CCSVI debate, forum users tried to validate the CCSVI hypothesis against their knowledge about MS in general (eg, by citing epidemiological facts such as the unequal sex ratio in MS) and their own illness experiences, such as symptom improvement by certain yoga techniques that are claimed to alter blood flow. In parallel to this embedding of CCSVI into the existing knowledge, users began to construct new experience-based knowledge about CCSVI. Two months after the first CCSVI-related posting, they began to publish first- or secondhand results of diagnostic and therapeutic procedures in the forum. In addition to this, they created hyperlinks to YouTube videos from MS patients from other parts of the world, which showed the CCSVI treatment and its positive outcome and also boosted the amount of experience-based knowledge about CCSVI when the evidence-based knowledge of the scientific community remained absolutely static without any new empirical results.

Collective Symbols

Many contributions to the CCSVI debate were colored by emotions, as could be observed from the frequent use of collective symbols, such as the German figure of speech "Halbgott in Weiß" ("demigod in white") used for degrading medical experts (see Textbox 1). Another example was the repeated use of "Bahnhof" ("railroad station"), a metaphor with connotations of "getting lost" or "being left behind."

We analyzed one figure of speech to reveal the underlying images from which the users constructed the picture of CCSVI and MS. The German expression "eine neue Sau durchs Dorf treiben" literally translates as "to chase a new sow through the village" and means to make a big fuss about something new, with a clearly negative connotation, and is often associated with the feeling that it distracts the audience's attention from the topic that really matters. In the context of CCSVI, the expression was always used to describe exactly this; the course of the CCSVI debate was sensed as familiar and repetitive, such as many other etiologic or therapeutic breakthroughs that were unable to keep the promise to heal MS. For some figures of speech, English-German matches exist. For example, "the early bird catches the worm" can be literally translated into German "Der frühe Vogel fängt den Wurm." In our case, it would be impossible to translate the figure of speech literally. Therefore, we give the original German quotation and its English denotation:

> *...In ein paar Monaten wird die Sau, die jetzt noch durchs MS-Dorf getrieben wird, tot zusammenbrechen, wie alle Wunderkur-Säue zuvor. Als alte MS-Hasen haben wir schon viele Säue verrecken sehen...[...few months from now, this final cure will all be revealed to be much ado about nothing, as were all the other final wonder-cures before. We've tried so many final cures before, we who've had MS for such a long time...]* [User 3, posted 1:08 pm October 17, 2009]
>
> *Stammzellen waren doch gestern. Heute sind's Krampfadern im Oberstübchen. Und morgen ne neue Sau, die durchs Dorf getrieben wird. Man darf also gespannt sein...[Stem cells are yesterday's news, aren't they. Today it's varicose veins in the belfry. And tomorrow it will be a new final cure. We're all absolutely holding our breath...]* [User 4, posted 10:46 pm August 16, 2009]
>
> *...Die "Krampfadern-im-Gehirn-Hypothese" kommt alle paar Jahre wieder, wenn wieder ein neues Publikum herangewachsen ist. Mal sehen, wie lang es diesmal dauert, bis die Sau sich durchs Dorf müde*

XSL•FO
RenderX

> *gerannt hat...[...The hypothesis of varicose veins in the belfry is repeated every few years when there's been a generation change in the auditorium. It'll be thrilling to watch how long it takes this time round until the bluff on the new final cure is called...]* [User 5, posted 10:41 am July 29, 2009]

The figure of speech portrays a certain emotional tableau of being disenchanted, hopeless, and being tired from past disappointments:

> *...Herzchen, die venöse Stau-Sau, die vielerorts getrieben wurde, ist hier schon durch. Wir haben sie gesehen, wahrgenommen, beklatscht, gewogen, für zu leicht befunden und in den nächsten Ort gejagt. Wenn Du Dich beeilst, dann kann es sein, dass Du sie noch einholst...Viel Glück. [...honey, the varicose-belfry apparition that's been reported from many locations has been through here already. We've seen it, noticed it, applauded it, finally weighed it and found it wanting. Then we chased it to the next town. If you make haste, you might still be catching up with it...Good luck.]* [User 6, posted 7:34 pm September 3, 2009]

Additionally, this enables the poster to distinguish him- or herself from those who do not notice that they are being messed with:

> *...Bei aller Verzweiflung, die man als MS-Kranker so hat: man muss nicht jeder Sau hinterherrennen, die grade mal wieder durchs Dorf getrieben wird...[...Even considering the heights of desperation that one experiences as an MS patient: you don't need to follow every self-proclaimed savior...]* [User 7, posted 12:37 pm October 17, 2009]

Characterization of Discourse Positions

We detected 4 different positions from which users participated in the CCSVI discussion and evaluated the associated incidents: "hostile," "frustrated," "wait and see," and "enthusiastic." These positions were not exclusive, meaning that there was a possibility to switch from one position to another: The hostile and the frustrated position formed a counterpart to the wait-and-see and enthusiastic positions; switching within these 2 groups, but not between them, seemed to be possible.

The positions differed in their orientation toward or against evidence-based and experience-based knowledge as reflected by the discourse strands described previously. The need for evidence-based information to assess CCSVI in its diagnostic and therapeutic value seemed at first glance to be accepted by most users. However, the argument that only prospective studies conducted by high-class research institutes would be able to produce reliable evidence-based information was disputed. This objection was not the result of a negative attitude toward research in general. Instead, the time that research needed to produce evidence-based knowledge was considered a price not every patient could afford to pay. Thus, the trust in evidence-based knowledge and the time pressure perceived simultaneously caused some inconsistency that was associated with negative feelings toward scientific research and scientists.

During lively debates of the significance of individual experiences or the trustworthiness of scientific information, the discussion became often highly emotional. These feelings colored the arguments or were directly verbalized. At some points in the discussion, the emotional coloring developed into a subjectivity, which often manifested itself in insults against others. A conspiracy theory seemed to exist on both sides: The pro-CCSVI side contested "Big Pharma" to the point of felonies like murder. However, the anti-CCSVI side also doubted the motives of CCSVI-promoting doctors and scientists. Economic interests were the main argument of both sides. Emotionality or subjectivity seemed to mask a deep personal involvement.

Considering the collective symbols, the figure of speech "to chase a new sow through the village" emblematically portrayed the discourse positions that were hostile to the enthusiastic position: the posters were considered disenchanted but wise in contrast to those who were enthusiastic but foolish. Table 1 shows the 4 discourse positions arranged according to the dimensions "trust in evidence-based knowledge," "trust in experience-based knowledge," and "subjectivity."

Table 1. Characterization of discourse positions during the chronic cerebrospinal venous insufficiency debate in the Deutsche Multiple Sklerose Gesellschaft (German Multiple Sclerosis Society) patient forum.

Dimension	Position			
	Hostile	Frustrated	Wait & see	Enthusiastic
Trust in evidence-based knowledge	Low	Low	Moderate	Moderate
Trust in experience-based knowledge	Low	Moderate	Moderate	High
Subjectivity	High	Low	Low	High

Discussion

In the discourse analysis of the CCSVI discussion in a German MS patient forum, we tried to reveal how patients reconcile the controversial scientific CCSVI debate with their own illness experience. The users heatedly debated whether scientific results or experiences of other patients had value for their own opinion-making about CCSVI. By determining the relation between relevant subtopics, we could generalize 2 discourse strands: (1) the "downfall of the professional knowledge providers" and (2) the "rise of the nonprofessional treasure trove of experience." The discourse strands indicated that the discussion moved away from the question whether scientific or experiential knowledge had more evidentiary value. Rather, the question whom to trust (ie, scientists, fellow sufferers, or no one at all) was of fundamental significance. Four discourse

XSL•FO
RenderX

positions could be identified by arranging them to the dimensions of "trust in evidence-based knowledge," "trust in experience-based knowledge," and "subjectivity." The emotional character of contributions was manifested by the use of popular rhetoric that seemed to mask a deep personal involvement.

At first glance, the lay discourse about the evidentiary value of experiential knowledge and the strong personal involvement may be interpreted as a misunderstanding of scientific methods, an intellectual incapacity to understand research practice, or even as irrationality. But this would be misjudging the discourses in patient online forums. Although most users were aware of the different positions in the scientific debate, at the same time they felt trapped in a conflictual relation between scientific and experiential knowledge, often associated with negative feelings of being cheated, left alone, and without control of their illness and their life. For that reason, the lay discourse in patient online forums should be interpreted as the result of the psychological efforts patients make to solve the conflictual tension arising from contradictory information.

The Tension Between Scientific Evidence and Personal Experience

The amount of and access to health information on the Internet is growing exponentially and numerous studies have investigated how well laypeople can assess and evaluate the quality of this information. These studies showed that laypeople only infrequently check the source and date of health information [18], can be misguided by search machines [19], and often misunderstand clinical concepts and aims of clinical research studies [20]. However, in our analysis of the CCSVI debate, the users of the MS forum sophisticatedly discussed the advantages and disadvantages of scientific research and the quality criteria of clinical studies, and they also searched for latest results of clinical research in scientific databases. At this point, the ideal of the "expert patient"—a term appearing for the first time in a report presented to the UK Parliament in 1999 as a "healthy citizen" initiative to help deal with chronic illness [21]—seems to come true. In addition to pursuing scientific knowledge, the users have gained experiential knowledge by (1) making various personal experiences by themselves as patients suffering from MS and (2) listening to the experiences of other patients. We could reveal that the users showed both scientific and experiential knowledge in the CCSVI discussion at different levels of elaboration and deliberation.

However, we witness a tension that emerges between these 2 different forms of knowledge. This tension can also be observed when health care professionals feel confronted with scientific evidence that disputed the medical practices they had been used to for years [22,23]. Interestingly, since the beginning when evidence-based medicine (EBM) was implemented as the new paradigm of health care, skeptical voices have discussed the problems resulting from its one-dimensional interpretation of evidence as scientific evidence [24].

Compared to this epistemological debate about EBM, our analysis of the CCSVI discussion in a patient online forum showed similar reactions using a somewhat different language. The result was the same (ie, a conflict between scientific and experiential knowledge). Although several users called for more time to give scientific proof a chance, others were immediately enthusiastic about the promises of the liberation treatment. The users seemed to sense an obligation to choose between the 2 forms of knowledge. Both parties claimed the higher evidentiary value of their knowledge.

The competition between these 2 forms of knowledge has a long tradition, with formally acquired knowledge—typified by objective science—being valued and naturalistic knowledge—typified by subjective experience—being devalued. To paraphrase Peter Storkerson, formal methods of knowledge acquisition are equated with rigor and validity and, consequently, knowledge derived from the application of such methods is valued per se as the gold standard of evidence [25]. In contrast, experiential knowledge is associated with unconscious, nonconscious, or implicit thinking that does not involve explicit, expressible, analyzable theoretical systems of knowledge. However, the users, or at least some of them, ascribed evidentiary value to experiential knowledge and to scientific knowledge. This could be interpreted as a kind of justification to choose the party of experiential knowledge. This evidence could be called "experiential evidence."

The conflictual tension between scientific and experiential evidence can be interpreted as "cognitive dissonance" [26], a discrepancy between action (ie, a forced choice between 2 alternatives) and attitude (ie, judging the 2 alternatives as being of the same value). Experimental social psychology has shown that people adjust their attitudes to support their decision by increasing their preference for the selected option, decreasing their preference for the rejected option, or both. This adjustment or rationalization is motivated by the urge to reduce the cognitive dissonance [27]. The theory of cognitive dissonance has been proven useful also in the context of health care research, for example, to examine patient behavior and emotional state after decisions concerning diagnostic or therapeutic procedures [27-29]. During the CCSVI debate, users might feel forced to choose between 2 alternatives of the same value on different levels: belief in scientific versus experiential evidence and also intellectual skepticism versus desperate hope.

A Matter of Trustworthiness

While discussing facts about and experiences with CCSVI in the forum, users were concerned about the quality of the information. They doubted that the DMSG was really neutral and patient-oriented; its impersonal stakeholders were assumed to be opportunistic and loyal to "Big Pharma." This reservation was also expressed toward scientists. They were perceived and portrayed as controlled by the pharmaceutical industry and driven by economic and career-related motives. In parallel to the devaluation of the DMSG and the MS research community, some other players were valorized by crediting them with trustworthiness. An example is Paolo Zamboni, the leader of the CCSVI movement. His trustworthiness was confirmed for the forum users because his wife has MS—this message was frequently communicated in the forum. This personal involvement resulted in a presumedly altruistic motivation.

This process of valorizing and devaluing different sources of knowledge is again in accordance with results of social psychology and the theory of cognitive dissonance. One way

XSL•FO
RenderX

to reduce cognitive dissonance is to challenge the source of the conflicting information (ie, to challenge its trustworthiness). Valorizing Zamboni by crediting him with trustworthiness was enhanced by the positive media presentation of Zamboni as the savior of his own wife and the liberation treatment like a miracle cure [30].

Disenchanted but Wise Versus Enthusiastic but Foolish

The different argumentation strategies and attitudes were revealed as the basis for the construction of 4 discourse positions in the CCSVI forum debate: the positions "hostile" and "frustrated" were contrasted with "wait and see" and "enthusiastic." These positions appear at first glance to be simple orientations or prejudices to uncritically adopt or reject new ideas. However, a more detailed examination showed that these argumentation strategies and attitudes were means to overcome the conflict between existing scientific results and experiential knowledge (ie, to reduce the cognitive dissonance). To reduce the cognitive dissonance in the CCSVI debate seemed to reflect a core experience in the course of the patients' illness experience: either you choose to distrust every new therapy and hypothesis to avoid being disappointed yet again or you choose to trust every new therapy and hypothesis in order not to miss the opportunity to be healed. The cognitive dissonance was perfectly symbolized by the often-used figure of speech "eine Sau durch das Dorf treiben" (to chase a sow through the village) that describes the situation of MS patients as either disenchanted but wise or enthusiastic but foolish.

There seems to be a relation between the 2 discourse positions with high subjectivity (ie "hostile" and "enthusiastic") and the level of personal involvement. We did not have any valid personal information about the contributors of the CCSVI discussion in the DMSG patient online forum (eg, age, sex, MS type). Statements about concrete relations between personal experiences as MS patient and the content and form of the contribution to the discussion would be highly speculative. However, reading between the lines (ie, when users casually talked about experiences they had as MS patients with the health care system and biomedical research) indicated different illness biographies. We sensed these illness biographies were dichotomized: biographies with a more benign course of the disease and positive experiences with academic medicine or biographies with a more devastating course and rather disappointing experiences with medicine and biomedical research. It might be speculated that the latter leads to a deeper personal involvement.

Strengths and Limitations

There are many articles discussing the developments in the context of the CCSVI debate, mainly comments or editorials [7,9]. However, articles delivering empirical results are rare. In addition to our own exploratory analysis [13], to our knowledge there is only one study that empirically investigated the CCSVI debate in the patient community with methods of the social sciences. The authors showed that patient experiences published as YouTube videos may replace evidence-based scientific information or create a hybrid of personal experience plus medical knowledge [8]. Although this analysis classified videos by their thematic content, our discourse analysis adds another piece of the puzzle of better understanding how patients try to reconcile the discrepancies between different forms of evidence.

Our study has some limitations. When using data that naturally emerged versus data produced for study purposes, such as survey data, the contributors are unlikely to be a representative sample of the population being analyzed. It is very likely that our sample of German MS patients consisted of more patients with a primary-progressive MS (PPMS) course than is true for the German MS population as a whole. Patients with PPMS are more often therapy refractory and have a fatal course with a higher grade of disability in a shorter period of time than patients with a relapsing-remitting MS [31]. The relation between a higher grade of disability or a worse course of disease and a more active commitment in patient forums and other social media has been described in the literature [32]. Patients with PPMS may more often have reservations against the health care system, biomedical science, and pharmaceutical industry, so that the CCSVI debate in our forum might be biased to a more negative course. Second, the analysis was based on only one online forum for German-speaking MS patients. Differences in culture, health care systems, and available treatments may influence the course of the CCSVI debate.

It would have been interesting to analyze in which way the discussion about CCSVI developed in the forum when the users had to face scientific evidence against the hypothesis and—perhaps more important—the liberation treatment. However, our main interest was to elucidate what happens in the lay discourse in ambivalent scientific situations. Ambivalence could arise when there are conflicting scientific results about one issue (eg, such as the case of breast cancer) [33]. Another example is the situation when a new hypothesis about a disease and its treatment spreads from the scientific community into the patient and lay community at a point of time when large clinical studies providing scientific evidence are still lacking. This was the case of CCSVI in the first 18 months after the initial Zamboni publication. Therefore, we decided to focus our analysis on this period of time.

Conclusions

We reconstructed a lay discourse of the evidentiary value of knowledge by critical discourse analysis of the CCSVI discussion in a patient forum and explained the development of the discourse with the theory of cognitive dissonance. This explanation puts the "expert patient" as proclaimed by health politics and research in question if "being an expert" means to be able to deliberate in cold blood the advantages and disadvantages of scientific evidence versus personal experience or to objectively evaluate conflicting results of scientific research. "Being a patient" means to be personally and emotionally affected, always "at risk of clutching at any straw." A healthy person without a family history of cancer balancing the advantages and disadvantages of taking part in the screening program is in a completely different situation than someone who is already handicapped by an illness and must fear losing control over his or her life. The need for cognitive dissonance reduction in this situation is so complex and urgent that it would be an unacceptable simplification to interpret the contributions of the forum users as a sign of irrationality and intellectual

XSL•FO
RenderX

incapacity of understanding scientific evidence or to blame them for trusting evidence that is based on experience.

Another unacceptable simplification might be the demand for science and scientists to regain lost credibility. The devaluation of science and scientists is probably not the result of the medial presentation of frauds and misconduct or the inability of scientists to explain their methods, results, and conclusions to the public. It is perhaps the consequence of patients' drive to reduce the cognitive dissonance that results from the conflictual tension between scientific and experiential knowledge. However, the role of the media in the CCSVI debate, particularly in Canada, was a good example of how the press presents new and sometimes absurd scientific ideas as "breakthrough" and "new hope" [34]. It is very likely that the media presentation has aggravated the uncomfortable feeling of cognitive dissonance in MS patients and will aggravate the dissonance between hope and skepticism for many other patients in the future.

Acknowledgments

One of the authors (JK) was funded by the research program of the University Medical Center Göttingen. This funding enabled her to establish the research group consisting of all other coauthors.

Conflicts of Interest

None declared.

References

1. 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]
2. Ziebland S, Wyke S. Health and illness in a connected world: how might sharing experiences on the internet affect people's health? Milbank Q 2012 Jun;90(2):219-249 [FREE Full text] [doi: 10.1111/j.1468-0009.2012.00662.x] [Medline: 22709387]
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. Zamboni P, Galeotti R, Menegatti E, Malagoni AM, Tacconi G, Dall'Ara S, et al. Chronic cerebrospinal venous insufficiency in patients with multiple sclerosis. J Neurol Neurosurg Psychiatry 2009 Apr;80(4):392-399 [FREE Full text] [doi: 10.1136/jnnp.2008.157164] [Medline: 19060024]
5. Vera C, Herr A, Mandato K, Englander M, Ginsburg L, Siskin GP. Internet-based social networking and its role in the evolution of chronic cerebrospinal venous insufficiency. Tech Vasc Interv Radiol 2012 Jun;15(2):153-157. [doi: 10.1053/j.tvir.2012.02.009] [Medline: 22640505]
6. Reekers JA, Lee MJ, Belli AM, Barkhof F. Cardiovascular and Interventional Radiological Society of Europe commentary on the treatment of chronic cerebrospinal venous insufficiency. Cardiovasc Intervent Radiol 2011 Feb;34(1):1-2 [FREE Full text] [doi: 10.1007/s00270-010-0050-5] [Medline: 21136256]
7. Fragoso YD. The internet racing ahead of the scientific evidence: the case of "liberation treatment" for multiple sclerosis. Arq Neuro-Psiquiatr 2011 Jun;69(3):525-527. [doi: 10.1590/S0004-282X2011000400022]
8. Mazanderani F, O'Neill B, Powell J. "People power" or "pester power"? YouTube as a forum for the generation of evidence and patient advocacy. Patient Educ Couns 2013 Dec;93(3):420-425 [FREE Full text] [doi: 10.1016/j.pec.2013.06.006] [Medline: 23830239]
9. Pullman D, Zarzeczny A, Picard A. Media, politics and science policy: MS and evidence from the CCSVI trenches. BMC Med Ethics 2013;14:6 [FREE Full text] [doi: 10.1186/1472-6939-14-6] [Medline: 23402260]
10. Reekers J. A swan song for CCSVI. Cardiovasc Intervent Radiol 2014 Apr;37(2):287-288. [doi: 10.1007/s00270-013-0833-6] [Medline: 24402646]
11. Ghezzi A. Funding CCSVI research is/was a waste of valuable time, money and intellectual energy: yes. Mult Scler 2013 Jun;19(7):855-857. [doi: 10.1177/1352458513479825] [Medline: 23712524]
12. Chafe R, Born KB, Slutsky AS, Laupacis A. The rise of people power. Nature 2011 Apr 28;472(7344):410-411. [doi: 10.1038/472410a] [Medline: 21525907]
13. Jäger S. Discourse and knowledge: theoretical and methodological aspects of critical discourse analysis. In: Wodak R, Meyer M, editors. Methods of Discourse Analysis. London: Sage Publications Ltd; 2001:32-62.
14. 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 Jan;16(1):e10 [FREE Full text] [doi: 10.2196/jmir.2875] [Medline: 24425598]
15. Anonymous. Patientenforum der Deutschen Multiple Sklerose Gemeinschaft. Der überstrapazierte Experte URL: http://dmsg.de/multiple-sklerose-forum/index.php?kategorie=forum&tnr=8&mnr=87435&aktpos_nav=0&keywords=Der%20%FCberstrapazierte%20Experte [WebCite Cache ID 6UOZnNlle]

XSL•FO
RenderX

16. Bond CS, Ahmed OH, Hind M, Thomas B, Hewitt-Taylor J. The conceptual and practical ethical dilemmas of using health discussion board posts as research data. J Med Internet Res 2013 Jun;15(6):e112 [FREE Full text] [doi: 10.2196/jmir.2435] [Medline: 23748147]
17. Schelling F. Damaging venous reflux into the skull or spine: relevance to multiple sclerosis. Med Hypotheses 1986 Oct;21(2):141-148. [Medline: 3641027]
18. Fox S. Online health research is widespread, but few check the source and date. MedGenMed 2007;9(1):30. [Medline: 17940593]
19. Allam A, Schulz PJ, Nakamoto K. The impact of search engine selection and sorting criteria on vaccination beliefs and attitudes: two experiments manipulating Google output. J Med Internet Res 2014 Apr;16(4):e100 [FREE Full text] [doi: 10.2196/jmir.2642] [Medline: 24694866]
20. Keselman A, Smith CA. A classification of errors in lay comprehension of medical documents. J Biomed Inform 2012 Dec;45(6):1151-1163 [FREE Full text] [doi: 10.1016/j.jbi.2012.07.012] [Medline: 22925723]
21. Department of Health. Saving Lives: Our Healthier Nation. London: Stationery Office; 1999.
22. Brown K, Sevdalis N. Lay vaccination narratives on the web: are they worth worrying about? Med Decis Making 2011 Sep;31(5):707-709. [doi: 10.1177/0272989X11419664] [Medline: 21921149]
23. Orr RD. Vertebroplasty, cognitive dissonance, and evidence-based medicine: what do we do when the 'evidence' says we are wrong? Cleve Clin J Med 2010 Jan;77(1):8-11 [FREE Full text] [doi: 10.3949/ccjm.77a.09146] [Medline: 20048024]
24. Buetow S, Kenealy T. Evidence-based medicine: the need for a new definition. J Eval Clin Pract 2000 May;6(2):85-92. [Medline: 10970002]
25. Storkerson P. Naturalistic cognition: A reason paradigm for human-centered design. J Res Pract 2010;6(2):2015-2003 [FREE Full text]
26. Festinger L. A Theory of Cognitive Dissonance. Stanford: Stanford University Press; 1957.
27. Elliot AJ, Devine PG. On the motivational nature of cognitive dissonance: dissonance as psychological discomfort. J Pers Soc Psychol 1994;67(3):382-394. [doi: 10.1037/0022-3514.67.3.382]
28. Gosling P, Denizeau M, Oberlé D. Denial of responsibility: a new mode of dissonance reduction. J Pers Soc Psychol 2006 May;90(5):722-733. [doi: 10.1037/0022-3514.90.5.722] [Medline: 16737370]
29. Steckelberg A, Kasper J, Muhlhauser I. Selective information seeking: can consumers' avoidance of evidence-based information on colorectal cancer screening be explained by the theory of cognitive dissonance? Ger Med Sci 2007;5:Doc05.
30. Laupacis A, Slutsky AS. Endovascular treatment for multiple sclerosis: the intersection of science, policy and the public. Open Med 2010;4(4):e197-e199 [FREE Full text] [Medline: 21687341]
31. Confavreux C, Vukusic S. The clinical course of multiple sclerosis. Handb Clin Neurol 2014;122:343-369. [doi: 10.1016/B978-0-444-52001-2.00014-5] [Medline: 24507525]
32. Han JY, Hou J, Kim E, Gustafson DH. Lurking as an active participation process: a longitudinal investigation of engagement with an online cancer support group. Health Commun 2014 Dec;29(9):911-923 [FREE Full text] [doi: 10.1080/10410236.2013.816911] [Medline: 24345206]
33. Maschke KJ. When evidence is contested. Hastings Cent Rep 2014;44(3):1p following 48. [doi: 10.1002/hast.314] [Medline: 24821253]
34. Nelkin D. An uneasy relationship: the tensions between medicine and the media. The Lancet 1996 Jun;347(9015):1600-1603. [doi: 10.1016/S0140-6736(96)91081-8]

Abbreviations

CCSVI: chronic cerebrospinal venous insufficiency
DSMG: Deutsche Multiple Sklerose Gesellschaft (German MS Society)
EBM: evidence-based medicine
MS: multiple sclerosis
PPMS: primary-progressive multiple sclerosis

Edited by G Eysenbach; submitted 04.12.14; peer-reviewed by F Mazanderani, A Allam; comments to author 26.02.15; revised version received 02.04.15; accepted 19.04.15; published 01.07.15.

Please cite as:
Koschack J, Weibezahl L, Friede T, Himmel W, Makedonski P, Grabowski J
Scientific Versus Experiential Evidence: Discourse Analysis of the Chronic Cerebrospinal Venous Insufficiency Debate in a Multiple Sclerosis Forum
J Med Internet Res 2015;17(7):e159
URL: http://www.jmir.org/2015/7/e159/
doi:10.2196/jmir.4103
PMID:26133525

XSL•FO
RenderX

©Janka Koschack, Lara Weibezahl, Tim Friede, Wolfgang Himmel, Philip Makedonski, Jens Grabowski. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 01.07.2015. 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

Twitter Social Media is an Effective Tool for Breast Cancer Patient Education and Support: Patient-Reported Outcomes by Survey

Deanna J Attai[1*], MD; Michael S Cowher[2*], MD; Mohammed Al-Hamadani[3*], MBChB, MPH; Jody M Schoger[4*]; Alicia C Staley[5*]; Jeffrey Landercasper[3,6*], MD

[1]UCLA Health Burbank Breast Care, Burbank, CA, United States

[2]Allegheny Health Network, Pittsburgh, PA, United States

[3]Gundersen Health System, Department of Medical Research, La Crosse, WI, United States

[4]Women With Cancer, The Woodlands, TX, United States

[5]Akari Healthcare, Charlestown, MA, United States

[6]Gundersen Health System, Norma J. Vinger Center for Breast Care, La Crosse, WI, United States

[*]all authors contributed equally

Corresponding Author:
Deanna J Attai, MD
UCLA Health Burbank Breast Care
191 S. Buena Vista Street
Suite 415
Burbank, CA, 91505
United States
Phone: 1 818 333 2555
Fax: 1 818 333 2559
Email: dattai@mednet.ucla.edu

Abstract

Background: Despite reported benefits, many women do not attend breast cancer support groups. Abundant online resources for support exist, but information regarding the effectiveness of participation is lacking. We report the results of a Twitter breast cancer support community participant survey.

Objective: The aim was to determine the effectiveness of social media as a tool for breast cancer patient education and decreasing anxiety.

Methods: The Breast Cancer Social Media Twitter support community (#BCSM) began in July 2011. Institutional review board approval with a waiver of informed consent was obtained for a deidentified survey that was posted for 2 weeks on Twitter and on the #BCSM blog and Facebook page.

Results: There were 206 respondents to the survey. In all, 92.7% (191/206) were female. Respondents reported increased knowledge about breast cancer in the following domains: overall knowledge (80.9%, 153/189), survivorship (85.7%, 162/189), metastatic breast cancer (79.4%, 150/189), cancer types and biology (70.9%, 134/189), clinical trials and research (66.1%, 125/189), treatment options (55.6%, 105/189), breast imaging (56.6%, 107/189), genetic testing and risk assessment (53.9%, 102/189), and radiotherapy (43.4%, 82/189). Participation led 31.2% (59/189) to seek a second opinion or bring additional information to the attention of their treatment team and 71.9% (136/189) reported plans to increase their outreach and advocacy efforts as a result of participation. Levels of reported anxiety before and after participation were analyzed: 29 of 43 (67%) patients who initially reported “high or extreme” anxiety reported “low or no” anxiety after participation (P<.001). Also, no patients initially reporting low or no anxiety before participation reported an increase to high or extreme anxiety after participation.

Conclusions: This study demonstrates that breast cancer patients’ perceived knowledge increases and their anxiety decreases by participation in a Twitter social media support group.

(J Med Internet Res 2015;17(7):e188) doi:10.2196/jmir.4721

KEYWORDS

breast cancer; education; social support; social media; patient outcome assessment

XSL•FO
RenderX

Introduction

Reporting that United States cancer care was "in crisis," the Institute of Medicine highlighted several critical areas for improvement in 2013 including the development of better methods for patient education, communication, shared decision making, and support [1]. They recommended 10 distinct goals, 3 of which are relevant to social media: providing patients with more "understandable information," psychosocial support, and to collect more patient-reported outcomes [1].

Traditional on-campus patient support groups are valuable and have been recommended by the National Accreditation Program for Breast Centers and other organizations in the United States [2,3]. However, traditional support groups may not be available for all patients. Patients may not participate because of inconvenient meeting times, transportation and child care issues, perception that their needs are not being specifically addressed, or reluctance to share their feelings or stories in public [4,5]. As an alternative, many women utilize various Internet and social media resources for medical information, advice, and support, such as blogs, chat groups, Facebook, and Twitter [4-7]. Social media is inherently bidirectional, interactive, and patient-driven in contrast with older models of health care education and decision making that are unidirectional and paternalistic.

Twitter is an online social networking service created in 2006 that enables users to send and read 140-character messages or "tweets." As of December 2014, Twitter had more than 280 million monthly active users [8]. Twitter has increasingly been embraced by patients as a way to share information and connect with other patients with similar concerns and conditions [7,9,10]. Disease-specific communities and chats have developed around "hashtags" (the # symbol followed by a word or acronym). Twitter chats now exist for patients with breast, lung, gynecologic, and pancreatic cancers, as well as multiple myeloma [11,12].

The Breast Cancer Social Media tweet chat (#BCSM) was initiated July 4, 2011, by 2 breast cancer survivors (authors JMS and ACS). Another author (DJA) became a comoderator in October 2011. The goal was to provide credible, evidence-based information and support for anyone affected by breast cancer. The chats occur on a weekly basis and cover all aspects of breast cancer screening, diagnosis, treatment, and survivorship. Specific medical advice is not provided and self-promotion or negativity toward any participant is actively discouraged.

Because of the growth and popularity of the #BCSM community, we sought to investigate the efficacy of Twitter social media to provide education and support to breast cancer patients. The study described in this paper is a pilot investigation assessing whether Twitter social media can provide education and psychosocial support to breast cancer patients.

Methods

A participant survey was developed to determine tweet chat participant characteristics and to assess the effectiveness of #BCSM for education and anxiety. Ten patient characteristics and 15 patient outcome domains were included. A 5-point Likert-type bipolar-scaled response was provided to survey participants for questions regarding whether their participation in #BCSM resulted in increased understanding for different domains of care and treatment. In addition, participants were queried for level of anxiety pre- and post-Twitter engagement, safety and comfort of participation in #BCSM, and motivation toward future advocacy and volunteer activities.

Surrogate measures of the impact of #BCSM in the Twitter community were obtained to include the number of tweets, the number of followers, and the product of multiplying these numbers, a measurement of the potential reach of #BCSM tweets [13].

A waiver of informed consent was obtained from the Gundersen Health System Institutional Review Board for the deidentified survey that was offered to Twitter participants. The survey link (Survey Monkey) was posted from April 14 to 24, 2014, on Twitter, the #BCSM blog, and the #BCSM Facebook page [14,15]. After survey closure, statistical analysis was performed to search for associations between patient characteristics and educational improvement for multiple domains of care. The majority of respondents were found to be in a single demographic group and some responders did not answer every survey question. This caused the sample sizes between different patient characteristics to be unbalanced, limiting statistical comparisons because of bias issues. Therefore, we limited analytic reporting to frequencies and percentages except for respondent-reported extreme or high anxiety before and after participation in the #BCSM tweet chats. The McNemar test was used for this comparison. Lastly, the data on overall Twitter participation with #BCSM and its impact based on number of impressions were determined.

Results

There were 206 survey responders. Respondent demographic and background information are presented in Table 1.

XSL•FO
RenderX

Table 1. Twitter participant characteristics (N=206).

Characteristic	n (%)[a]
Age (years)	
≤24	1 (0.5)
25-34	21 (10.2)
35-44	46 (22.3)
45-54	69 (33.5)
55-64	57 (27.6)
65-74	11 (5.3)
≥75	1 (0.5)
Sex	
Male	15 (7.2)
Female	191 (92.7)
Community population size	
<10,000	15 (7.3)
10,000-100,000	46 (22.3)
100,001-1,000,000	72 (34.9)
>1,000,000	73 (35.4)
Race/ethnicity	
White (includes Latino and Hispanic)	189 (91.7)
Black or African American	4 (1.9)
North, South, or Central American, Native Indian, or Alaskan native	1 (0.5)
Native Hawaiian or other Pacific Islander	2 (1.0)
Asian	5 (2.4)
Other	5 (2.4)
Highest level of education	
Primary school	0 (0.0)
Some high school, but no diploma	0 (0.0)
High school diploma (or GED)	1 (0.5)
Some college, but no degree	22 (10.7)
2-year college degree	11 (5.3)
4-year college degree	59 (28.6)
Graduate-level degree	111 (53.9)
None of the above	2 (0.9)
Previous or current treatment for breast cancer	
Yes	143 (69.4)
No	63 (30.5)
If not been treated for breast cancer, which category describes you	
Caregiver, family, friend, or spouse to breast cancer patient	17 (25.4)
Clinical health professional or researcher	25 (37.3)
Other	25 (37.3)
Length of participation in chats (months)	
<6	59 (28.6)
6-12	54 (26.2)

XSL•FO
RenderX

Characteristic	n (%)[a]
>12	93 (45.1)
Engagement with #BCSM community outside of Twitter	
Yes	141 (77.0)
No	42 (22.9)
Participation in in-person support groups	
Yes	79 (43.2)
No	**104 (56.8)**

[a] Not all response categories add up to 206 total survey responders because not all participants answered every question.

A total of 92.7% (191/206) of respondents were female and 69.4% (143/206) of respondents were breast cancer patients. Other respondents included family, friends, advocates, surgeons, medical oncologists, radiation oncologists, clinical psychiatrists, genetic counselors, and physical therapists. Respondent-reported changes in level of understanding and education in multiple domains of care after tweet chat participation were assessed. Survey participants were asked whether participation "provided a safe and welcoming forum for support and education." Of 183 respondents to this question, 116 (63.4%) "strongly agreed," 44 (24.0%) "somewhat agreed," 12 (6.5%) were "neutral," 8 (4.4%) "somewhat disagreed," and 3 (1.6%) "strongly disagreed." As a result of #BCSM, 52 of 183 responders (28.4%) reported subsequent volunteer efforts. For analysis of respondent reports of anxiety, survey responders reporting on less than 50% of survey questions were excluded. The 189 remaining responders were analyzed for their recall levels of reported high/extreme anxiety before and after participation in the #BCSM tweet chats. We found a significant decrease in the proportion of respondents with extreme/high anxiety level from 43 of 153 patients (28.1%) to 14 of 152 (9.2%, $P<.001$). Also, no respondents who initially reported "low or no" anxiety before participation reported an increase to "high or extreme anxiety" after participation. See Table 2 for other respondent answers.

Overall Twitter participation with #BCSM and its potential reach based on number of impressions is shown in Table 3 [13].

Table 2. Improvement in knowledge level after Twitter participation (N=206).

Knowledge domain	Response, n (%)[a]				
	Definitely yes	Somewhat yes	Not sure	Somewhat no	Definitely no
Breast cancer type[b]	66 (34.9)	68 (35.9)	22 (11.6)	16 (8.4)	17 (8.9)
Surgery and reconstruction	38 (20.1)	67 (35.4)	33 (17.5)	23 (12.2)	28 (14.8)
Radiation treatment	31 (16.4)	51 (26.9)	47 (24.9)	25 (13.2)	35 (18.5)
Breast imaging	39 (20.6)	68 (35.9)	33 (17.5)	21 (11.1)	28 (14.8)
Metastatic stage 4 cancer	95 (50.2)	55 (29.1)	17 (8.9)	9 (4.7)	13 (6.9)
Clinical trials	45 (24.3)	80 (42.3)	30 (15.9)	17 (8.9)	17 (8.9)
Genetic risk	36 (19.0)	66 (34.9)	45 (23.8)	24 (12.6)	18 (9.5)
Survivorship[c]	99 (52.4)	63 (33.3)	11 (5.8)	8 (4.2)	8 (4.2)
Advocacy and fundraising	70 (37.0)	66 (34.9)	33 (17.5)	13 (6.9)	7 (3.7)
Healthy lifestyle[d]	40 (21.2)	64 (33.9)	44 (23.3)	18 (9.5)	23 (12.2)
My impact on others	76 (40.2)	70 (37.0)	25 (13.2)	9 (4.8)	9 (4.8)
Seek second opinion[e]	29 (15.3)	30 (15.9)	19 (10.1)	18 (9.5)	49 (25.9)
Overall education with "your" cancer[e]	31 (16.4)	84 (44.4)	42 (22.0)	0 (0.0)	0 (0.0)
Overall education with any breast cancer	58 (30.7)	95 (50.2)	35 (18.5)	1 (0.5)	0 (0.0)

[a] The number of respondents included in the denominator differs for each survey question because some respondents did not answer every question.

[b] Understanding of estrogen receptors (ER) and progesterone receptors (PR), HER2, triple-negative types and meaning.

[c] Understanding posttreatment follow-up side effects, lymphedema (arm swelling), cognitive impairment from chemotherapy, sexuality, grief, death, or other.

[d] Diet, exercise, lifestyle habits.

[e] Does not equal to 100% because the question was not applicable to the noncancer participants.

XSL•FO
RenderX

Table 3. Annual trends of Twitter participation with #BCSM.

Twitter participation	Year			
	2011	2012	2013	2014
Impressions,[a] n	101,263,199	309,657,740	295,718,132	343,586,925
Impressions per user, mean	48,359	71,963	35,853	24,220
Tweets,[b] n	28,275	69,505	84,614	85,972
Tweets per user, mean	13.5	16.2	10.3	6.1
Users,[c] n	2094	4303	8248	14,186

[a] The (number of tweets)*(number of followers), a measurement of the potential reach of a tweet.

[b] Tweet that went out tagged with #BCSM.

[c] Unique individuals who posted anything with #BCSM.

Discussion

Major health care policy stakeholders in the United States endorse a model of patient care that is not solely provider- or institution-directed [1,16-22]. These stakeholders recognize that differences may exist between providers, payers, and patients regarding what constitutes good care and how to measure it [23]. Consequently, they have developed tools and resources that promote and measure patient-centered care, such as the Patient Center for Outcomes Research, the Consumer Assessment of Healthcare Providers and Systems (CAHPS) surveys, and breast surgery patient surveys [18-21]. Unfortunately, the existing resources for measuring patient quality of life, such as CAHPS surveys, have limitations because they require significant infrastructure and financial support for their use. In contrast, social media has the ability to aid patient centeredness, report patient outcomes, and provide breast specialists with patient perceptions of gap in care with limited financial and information technology investment. Social media is user-friendly and popular for health care consumers as evidenced by its rapid growth. Moreover, because of the widespread use and potential inclusiveness of nearly all patient demographic groups in social media, it is fertile territory for future investigations regarding identification of gaps in care, along with its potential to measure the success of patient education and support. This study is an initial investigation into the ability of Twitter social media to improve breast cancer patient education and support. Our primary aim was to determine patient's perception of level of benefit of participation.

Only a few other reports of benefits of participation in an online social media-based support group have been published, although cancer survivors report the Internet as the second most important source for cancer information after their health care professional [6,24]. Online support groups have been shown to fill gaps in supportive care by meeting needs of some breast cancer survivors [25].

Approximately 20% to 30% of breast cancer survivors demonstrate measurable signs of anxiety and/or depression in the year following diagnosis [26,27]. Although the symptoms are most pronounced during the first year after diagnosis, up to 15% have depressive symptoms 5 years after diagnosis. Our study demonstrated an association between a reduction in breast cancer-perceived anxiety and participation in the #BCSM tweet chats. Other measures of success were also demonstrated. Twitter participants reported improved knowledge about multiple domains of care through the continuum of the cancer care timeline, inferring traditional educational resources may not have been sufficient. Despite this success, Twitter and other forms of social media cannot replace traditional office- and hospital-based resources for education and emotional support. In our survey, nearly 1 in 5 Twitter participants reported no improvement in education and 9% had persistent high anxiety despite #BCSM suggesting that Twitter participation was not sufficient to address education and anxiety needs in these patients. Therefore, the role of social media is to compliment but not replace current practice. It is reassuring that nearly 90% of Twitter followers reported the #BCSM chats as a "safe and comfortable" environment, given that more than half of participants reported no involvement in any other support group. A novel finding of this study is the observation that many responders to our survey were motivated by #BCSM to participate in advocacy or volunteer efforts.

There are many limitations to this pilot social media study. Our patient survey has not undergone formal reliability and validity testing and the survey format is subject to recall bias. We cannot determine if recall bias occurred because our survey was open for participation at only one snapshot in time. In addition, there was no control group of patients, with similar demographic characteristics, who had never participated in Twitter and were concurrently surveyed for their level of education and anxiety over a similar time period. Our survey cohort was homogenous, mostly white, and well educated. This lack of diversity prevents any statistical comparisons between demographic groups for Twitter's ability to educate or ameliorate anxiety. The lack of diversity also does not allow extrapolation of study results to other patient populations to include non-Twitter users who may have differences in their demographic profile and cancer status compared to Twitter users.

It is unclear if the respondents to the survey were truly representative of the #BCSM community at large. Given the public nature of Twitter, there is no way to fully assess the participant composition of any Twitter chat. Of the survey respondents, 45% reported that they observe the chats but do not participate. "Lurking" (observing but not actively participating) on social media forums is a well-described

XSL•FO
RenderX

phenomenon, although it has been reported that lurkers obtain the same benefits in terms of empowerment as those who actively participate [28,29].

Despite these limitations, our results represent an important first step toward development of more online patient education and support communities. An association between Twitter participation and improvements in patient self-reported knowledge and anxiety was identified. Compared to other cancer patients, those with breast cancer are more likely to seek online information about their condition [30]. Health care providers utilizing social media serve as patient advocates, assisting patient and family understanding of their disease and its treatment. Physician participation in online patient communities can help counter inaccurate and sometimes dangerous information. Other benefits of physician use of social media have been described [31].

Since the first #BCSM tweet chat on July 4, 2011, more than 160 chat hours have been logged. Chat activity, measured by the overall number of tweets, has increased each year as has the unique number of participants. Other cancer-related tweet chats have developed (Table 4), including sites for lung cancer, gynecologic cancers, multiple myeloma, and pancreatic cancer.

Table 4. Cancer-related tweet chat sites.

Cancer type	Hashtag (#)	Physician moderator(s)	Twitter handle
Breast	#BCSM	Dr Deanna J Attai	@DrAttai
Lung	#LCSM	Dr H Jack West	@JackWestMD
		Dr. David T. Cook	@UCD_ChestHealth
Gynecologic	#GYNCSM	Dr Don Dizon	@drdonsdizon
		Dr Merry Markham	@DrMarkham
		Dr Rick Boulay	@journeycancer
Multiple myeloma	#MMSM	Dr Mike Thompson	@mtmdphd
Pancreatic	#PANCSM	Dr Niraj Gusani	@NirajGusani
		Dr Mark Bloomston	@pancdoc
		Dr Matthew HG Katz	@mkatzmd
		Dr Diane Reidy-Lagunes	@DianeReidyLagun

A unique aspect of these chats is patient-physician involvement; in fact, most are comoderated by patients and physicians. The patients participating are seeking credible evidence-based information—something that may be lacking in other online forums. For physicians wary of the social media environment, #BCSM and the other cancer-related tweet chats provide an example of positive patient-physician engagement.

By surveying Twitter participants, we demonstrated proof of concept of improved patient knowledge regarding their disease-specific condition and management. Further investigations are warranted to explore its capability to provide increased patient knowledge, psychosocial support, and meaningful networking between patients and caregivers.

Acknowledgments

The authors acknowledge the contributions of Choua Vang for her assistance in manuscript preparation. Jeffrey Landercasper, Choua Vang, and Mohammed Al-Hamadani acknowledge the financial assistance of the Gundersen Medical Foundation and the Norma J Vinger Center for Breast Care, Gundersen Health System, La Crosse, Wisconsin.

Conflicts of Interest

Deanna J Attai, Michael S Cowher, Mohammed Al-Hamadani, Jody M Schoger, and Jeffrey Landercasper have no disclosures. Alicia C Staley has the following disclosures: CEO of Akari Health; ePharma: Clinical Trials Speaker, Honorarium, January 2014; Robert Wood Johnson Foundation, Speaker, Honorarium February 2014; and Livestrong Conference Speaker, Honorarium, May 2014.

References

1. Institute of Medicine. Delivering high-quality cancer care: charting a new course for a system in crisis. 2013 Sep 10. URL: http://iom.nationalacademies.org/Reports/2013/Delivering-High-Quality-Cancer-Care-Charting-a-New-Course-for-a-System-in-Crisis.aspx?utm_source=feedburner&utm_medium=feed&utm_campaign=Feed%3A+HomTopicHealthServicesCoverageAndAccess+(IOM+Topic%3A+Health+Services+Coverage+and+Access) [accessed 2015-04-28] [WebCite Cache ID 6Y8BhLNqx]

XSL•FO
RenderX

2. National Accreditation Program for Breast Centers. NAPBC Standards Manual 2014 Edition. Chicago, IL: American College of Surgeons; 2014. URL: https://www.facs.org/~/media/files/quality%20programs/napbc/2014%20napbc%20standards%20manual.ashx [accessed 2015-04-28] [WebCite Cache ID 6Y8GhjbXO]
3. Lepore SJ, Buzaglo JS, Lieberman MA, Golant M, Davey A. Standard versus prosocial online support groups for distressed breast cancer survivors: a randomized controlled trial. BMC Cancer 2011;11:379 [FREE Full text] [doi: 10.1186/1471-2407-11-379] [Medline: 21867502]
4. Ussher JM, Kirsten L, Butow P, Sandoval M. A qualitative analysis of reasons for leaving, or not attending, a cancer support group. Soc Work Health Care 2008;47(1):14-29. [Medline: 18956510]
5. Stephen J, Rojubally A, Macgregor K, McLeod D, Speca M, Taylor-Brown J, et al. Evaluation of CancerChatCanada: a program of online support for Canadians affected by cancer. Curr Oncol 2013 Feb;20(1):39-47 [FREE Full text] [doi: 10.3747/co.20.1210] [Medline: 23443892]
6. Hong Y, Peña-Purcell NC, Ory MG. Outcomes of online support and resources for cancer survivors: a systematic literature review. Patient Educ Couns 2012 Mar;86(3):288-296. [doi: 10.1016/j.pec.2011.06.014] [Medline: 21798685]
7. Tsuya A, Sugawara Y, Tanaka A, Narimatsu H. Do cancer patients tweet? Examining the twitter use of cancer patients in Japan. J Med Internet Res 2014;16(5):e137 [FREE Full text] [doi: 10.2196/jmir.3298] [Medline: 24867458]
8. Twitter. About Twitter URL: https://about.twitter.com/company [accessed 2015-04-28] [WebCite Cache ID 6Y8GrSho8]
9. Himelboim I, Han JY. Cancer talk on twitter: community structure and information sources in breast and prostate cancer social networks. J Health Commun 2014;19(2):210-225. [doi: 10.1080/10810730.2013.811321] [Medline: 24111482]
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. Katz M. Symplur. 2013 Nov 04. Hashtags in cancer care: embedding meaning in digital health URL: http://www.symplur.com/blog/hashtags-cancer-care-embedding-meaning-digital-health/ [accessed 2015-04-28] [WebCite Cache ID 6Y8GzWOrL]
12. Symplur. About cancer tag ontology URL: http://www.symplur.com/healthcare-hashtags/ontology/cancer/ [accessed 2015-04-27] [WebCite Cache ID 6Y8H5D6Sa]
13. Symplur. 2015. Potential "reach" of Twitter URL: http://www.symplur.com/healthcare-hashtags/bcsm/ [accessed 2015-06-22] [WebCite Cache ID 6ZTfjYcHR]
14. Breast Cancer Social Media. 2015. Welcome to the #BCSM Community URL: http://bcsmcommunity.org/ [accessed 2015-04-29] [WebCite Cache ID 6Y9ZIIv93]
15. Facebook. BCSM Community URL: https://www.facebook.com/BCSMCommunity [accessed 2015-04-29] [WebCite Cache ID 6Y9ZdjHp2]
16. Savitz L, Luther K. Patient-reported measures. Healthcare Executive 2015 Jan;30(1):74-77.
17. Greenberg CC, Kennedy GD. Advancing quality measurement to include the patient perspective. Ann Surg 2014 Jul;260(1):10-12. [doi: 10.1097/SLA.0000000000000730] [Medline: 24896145]
18. The National Quality Strategy. NQS tools URL: http://www.ahrq.gov/workingforquality/toolkit.htm [accessed 2015-04-28] [WebCite Cache ID 6Y8HNaTh5]
19. Sage J. Using S-CAHPS. Bull Am Coll Surg 2013 Aug;98(8):53-56. [Medline: 24205576]
20. Agency for Healthcare Research and Quality. Consumer Assessment of Healthcare Providers and Systems (CAHPS) surveys URL: https://cahps.ahrq.gov/ [accessed 2015-04-28] [WebCite Cache ID 6Y8HZ4PU6]
21. The American Society of Breast Surgeons. Breast cancer surgery patient survey to measure patient reported outcomes in the American Society of Breast Surgeons Mastery Program URL: https://www.breastsurgeons.org/new_layout/programs/mastery/index.php [accessed 2015-04-28] [WebCite Cache ID 6Y8HgvjFN]
22. Patient-Centered Outcomes Research Institute. 2015. About us URL: http://www.pcori.org/about-us [accessed 2015-04-28] [WebCite Cache ID 6Y8HsOSld]
23. Bingener J, Sloan J, Boughey JC. Patient quality of life: vitally important. Bull Am Coll Surg 2014 May;99(5):46-48. [Medline: 24868847]
24. Lepore SJ, Buzaglo JS, Lieberman MA, Golant M, Greener JR, Davey A. Comparing standard versus prosocial internet support groups for patients with breast cancer: a randomized controlled trial of the helper therapy principle. J Clin Oncol 2014 Dec 20;32(36):4081-4086. [doi: 10.1200/JCO.2014.57.0093] [Medline: 25403218]
25. Bender JL, Katz J, Ferris LE, Jadad AR. What is the role of online support from the perspective of facilitators of face-to-face support groups? A multi-method study of the use of breast cancer online communities. Patient Educ Couns 2013 Dec;93(3):472-479. [doi: 10.1016/j.pec.2013.07.009] [Medline: 23928354]
26. Björneklett HG, Lindemalm C, Rosenblad A, Ojutkangas M, Letocha H, Strang P, et al. A randomised controlled trial of support group intervention after breast cancer treatment: results on anxiety and depression. Acta Oncol 2012 Feb;51(2):198-207. [doi: 10.3109/0284186X.2011.610352] [Medline: 21923569]
27. Björneklett HG, Rosenblad A, Lindemalm C, Ojutkangas M, Letocha H, Strang P, et al. Long-term follow-up of a randomized study of support group intervention in women with primary breast cancer. J Psychosom Res 2013 Apr;74(4):346-353. [doi: 10.1016/j.jpsychores.2012.11.005] [Medline: 23497838]

XSL•FO
RenderX

28. 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;10(2):e18 [FREE Full text] [doi: 10.2196/jmir.992] [Medline: 18653442]
29. Han JY, Kim J, Yoon HJ, Shim M, McTavish FM, Gustafson DH. Social and psychological determinants of levels of engagement with an online breast cancer support group: posters, lurkers, and nonusers. J Health Commun 2012;17(3):356-371 [FREE Full text] [doi: 10.1080/10810730.2011.585696] [Medline: 22085215]
30. Nagler RH, Gray SW, Romantan A, Kelly BJ, DeMichele A, Armstrong K, et al. Differences in information seeking among breast, prostate, and colorectal cancer patients: results from a population-based survey. Patient Educ Couns 2010 Dec;81 Suppl:S54-S62 [FREE Full text] [doi: 10.1016/j.pec.2010.09.010] [Medline: 20934297]
31. Maisel M. Four reasons doctors should use social media. Breast Diseases: A Year Book Quarterly 2014;25(4):302-305. [doi: 10.1016/j.breastdis.2014.10.033]

Abbreviations

CAHPS: Consumer Assessment of Healthcare Providers and Systems

Edited by G Eysenbach; submitted 19.05.15; peer-reviewed by H Narimatsu, E Castro-Sánchez; comments to author 16.06.15; revised version received 22.06.15; accepted 08.07.15; published 30.07.15.

Please cite as:
Attai DJ, Cowher MS, Al-Hamadani M, Schoger JM, Staley AC, Landercasper J
Twitter Social Media is an Effective Tool for Breast Cancer Patient Education and Support: Patient-Reported Outcomes by Survey
J Med Internet Res 2015;17(7):e188
URL: http://www.jmir.org/2015/7/e188/
doi:10.2196/jmir.4721
PMID:26228234

©Deanna J Attai, Michael S Cowher, Mohammed Al-Hamadani, Jody M Schoger, Alicia C Staley, Jeffrey Landercasper. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 30.07.2015. 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

A Privacy Preservation Model for Health-Related Social Networking Sites

Jingquan Li[1], PhD, MS, BE (CS)

Health Informatics Research, College of Business, Texas A&M University-San Antonio, San Antonio, TX, United States

Corresponding Author:
Jingquan Li, PhD, MS, BE (CS)
Health Informatics Research
College of Business
Texas A&M University-San Antonio
One University Way
San Antonio, TX, 78224
United States
Phone: 1 2107842312
Fax: 1 2107842398
Email: jli@tamusa.tamus.edu

Abstract

The increasing use of social networking sites (SNS) in health care has resulted in a growing number of individuals posting personal health information online. These sites may disclose users' health information to many different individuals and organizations and mine it for a variety of commercial and research purposes, yet the revelation of personal health information to unauthorized individuals or entities brings a concomitant concern of greater risk for loss of privacy among users. Many users join multiple social networks for different purposes and enter personal and other specific information covering social, professional, and health domains into other websites. Integration of multiple online and real social networks makes the users vulnerable to unintentional and intentional security threats and misuse. This paper analyzes the privacy and security characteristics of leading health-related SNS. It presents a threat model and identifies the most important threats to users and SNS providers. Building on threat analysis and modeling, this paper presents a privacy preservation model that incorporates individual self-protection and privacy-by-design approaches and uses the model to develop principles and countermeasures to protect user privacy. This study paves the way for analysis and design of privacy-preserving mechanisms on health-related SNS.

(J Med Internet Res 2015;17(7):e168) doi:10.2196/jmir.3973

KEYWORDS

social networks; privacy; security; threat modeling; privacy preservation model; electronic health records; health care

Introduction

Health-related social networking sites (SNS) are websites that enable the connection of users and facilitate the exchange of health knowledge and information. Physicians can connect with their peers and collaborate on patient cases and other medical topics to improve health care delivery and patient outcomes at sites like Doximity. Patients with life-changing illnesses can find other patients like them, discuss and track medical conditions, and give and receive support at PatientsLikeMe. Before the advent of SNS, medical providers and pharmaceutical companies spread the word to encourage participation in wellness and disease management programs. Today, websites such as Inspire and DailyStrength provide users with the opportunity to share information and stories about healthy living, thereby supporting and inspiring others. The proliferation of these sites is building a new health-information technology business prophesied on the belief that the wisdom of crowds really is smarter than any one person, no matter how well researched the individual person.

However, users reveal vast amounts of personal health information on a health-related SNS. They may also join other social networks or websites and enter personal information and other specific information covering social, professional, and health domains into other websites. There are many possible ways that users' privacy can be compromised: data misuses, disclosures to intruders, accidental data releases, disclosures to third parties and apps, and user profiling across multiple social networks. A recent incident in which a major media monitoring firm improperly scraped personal data from PatientsLikeMe demonstrates significant privacy risks for online health information [1]. In the United States, health care providers

XSL•FO
RenderX

disclose patient information without patient authorization in violation of the Health Insurance Portability and Accountability Act (HIPAA), the Health Information Technology for Economic and Clinical Health (HITECH) Act, and/or state privacy laws and can be subject to fines and other penalties. In the age of Facebook and Twitter, however, many patients themselves volunteer to post their personally identifiable information (PII) and sensitive health information on multiple social networks.

While privacy concerns in social networks are well recognized by prior research [2-7], the literature on innovative privacy-preserving models and technical standards is quite limited. Based on expert opinions on the major privacy concerns, the effectiveness of possible solutions, and the requirements for developing privacy-preserving social network apps, Weiss [7] proposed a privacy threat model for data portability in social network apps. However, this work concentrated on privacy in the sense of visibility and transparence, that is, transparent and open privacy handling practices, and not so much on the privacy-preserving mechanisms that need to be developed. To address the privacy issues caused by the central SNS provider, such as data misuse and leakage risk, it has been proposed to decentralize social networking services [8-10]. However, to our best knowledge, current health-related SNS are predominantly logically centralized services and the underlying business model relies on access to the user-generated content, resulting in the impracticality of the decentralized SNS approach. There is a strong need to develop a privacy model that can protect user privacy in the complex social networking environment. This paper addresses this gap by identifying the most important threats to users and SNS providers and proposes a privacy preservation model to address the privacy challenges of health-related SNS. This paper first analyzes privacy concerns related to health-related SNS. It then develops a threat model and articulates some principal threats. Since current privacy solutions such as end-user license agreements and privacy settings are inadequate to address the threats, this paper presents a privacy preservation model that integrates both individual self-protection and privacy-by-design approaches and uses the model to develop principles and countermeasures to protect privacy.

Privacy Concerns Related to Social Networking Sites

To illustrate privacy issues with health-related SNS, we analyze the Inspire platform. Inspire may be considered an illustrative case of patient sites that offer a privacy policy and settings to address users' privacy concerns.

Inspire is an online health and wellness support community for patients and caregivers. Inspire is provided by ClinicaHealth, Inc., and is composed of more than 190 disease-specific communities. As of December 2014, Inspire has over 400,000 registered members and 700,000 unique visitors each month.

Inspire is free for individuals and non-profit patient advocacy associations [11]. Its business model largely depends on advertising revenue and partnerships with many third-party companies [12]. Inspire helps companies and researchers find likely clinical trial participants. Clinical trial sponsors pay Inspire for this service. Inspire also offers health-focused research services to commercial companies. User-generated content has high value for the companies to conduct secondary research and issues analysis. Furthermore, Inspire makes money from selling targeted advertising.

The products and services of Inspire are essentially user profiles and user-generated content. Inspire collects and stores three types of information from users: personal profile, user-generated content, and Web behavior information. At its registration page, Inspire asks a new user to provide certain personal information, including a functioning email address, postal code, gender, date of birth, user ID, and password. A user is also given an option to provide additional personal information to create an extended online profile [13]. User-generated content is all the information a user posts on the site or communicates with other users, including disease conditions, treatments, family history, and possibly personal information generated by the user. Web behavior information is information on how a user uses different features of the site collected through cookies. Inspire may combine this information with the profile [13].

Inspire strives to create a secure environment where users connect with each other around shared conditions and share relevant information about their health and the health of their loved ones. When people's personal information is involved, however, there are several privacy concerns. First, Inspire may reveal personal information to other users and outsiders. When a user registers at the site, the profile becomes visible to other users of the site and the profile may also be found by visitors of the site using Inspire's search functions. Although users can use the privacy settings to control access to their profiles, they may not have the knowledge and technical skills to understand the settings and change their own settings appropriately.

Second, Inspire has the right to use personal information for various purposes without user control. For example, it may use personal information to present targeted content, including advertising or requests either from ClinicaHealth or from a third party. Users have no control over the collection and use of personal information by Inspire and its affiliates. Inspire makes clear under its privacy policy: "ClinicaHealth may share your email address and profile information with the organizations that sponsor Inspire groups that you join" [13].

Third, although Inspire does not disclose a user's PII to third parties without consent, it may share health information with third parties on an aggregate or other basis that does not disclose user identity or contain PII [13]. However, concerns have been raised about the sufficiency of popular de-identification methodologies such as merely stripping names and addresses; data mining tools make it possible to reverse-engineer PII from weakly de-identified user information [1]. Furthermore, user-generated content, which may contain PII accidentally revealed by users, is open to the community, outsiders, and third parties.

Fourth, Inspire is an open community. Anyone with a valid email address can sign up for Inspire and then view the content on the site. This raises the problem of unauthorized access by unintended users. Inspire is also vulnerable to attacks from

XSL•FO
RenderX

malicious intruders, such as data scraping and social engineering attacks.

Table 1. Examples of health-related social networks and general social networks.

Social network	Description	Privacy practices
CarePages.com	CarePages is a community of people collaborating to share the challenges, hopes, and victories of anyone facing a life-altering health event.	Privacy settings include “Community”, “Friends & Family”, and “Invitation Only”; secondary use of personal information; CarePages combines personal information with the data received from third parties to target advertising [14].
CureTogether.com	CureTogether provides a service whereby patients and researchers come together to share information and find cures for chronic diseases.	Privacy settings include “Public”, “Research”, “Friends”, and “Private”; secondary use of personal information; disclosing de-identified information to third-party researchers [15].
DailyStrength.org	DailyStrength is a health network of people sharing advice, treatment experiences, and support.	Users and visitors can see any information users provide; secondary use of personal information; DailyStrength reserves the right to use and disclose de-identified information to third parties at its discretion [16].
Inspire.com	Inspire has mini social networks for different diseases and health conditions, each sponsored by health organizations.	Privacy settings include “Public”, “Members”, “Friends”, and “Private”; secondary use and disclosure by the SNS provider and its affiliates; sharing aggregate personal and health information with third parties [13].
PatientsLikeMe.com	PatientsLikeMe is a social network that enables people to share health experiences that can improve the lives of patients diagnosed with chronic diseases.	PatientsLikeMe provides two privacy levels “Public” and “Members”; secondary use of personal information; disclosing shared data to partners and other third parties for use in scientific research and market research [17].
Facebook.com	Facebook is a social network that enables users to create profiles, upload photos and videos, send messages, and communicate with friends, family, and colleagues.	Privacy settings include “Public”, “Friends”, “Only Me”, “Custom”, and “Close Friends”; secondary use of personal information; sharing non-personally identifiable information with advertising and analytics services and disclosing all information to other third parties [18].
Twitter.com	Twitter is a microblogging platform that enables users to send and read short 140-character messages called “tweets”.	Tweets can be “Public” or “Private”; a public user profile, login verification, and tweet location can be configured; secondary use of personal information; sharing personal information with its service providers and third parties [19].

These concerns are also intrinsic to other health-related SNS and general social networks such as Facebook and Twitter. The purpose and privacy characteristics of top patient sites and general social networks are summarized in Table 1 [14-19]. Furthermore, the security characteristics of these sites are vague. Without effective privacy controls, health-related SNS may disclose the information not only to their business partners but also to unintended individuals and entities. The concern is not just about data mining and marketing that could influence patients to seek drugs they do not need or to spend more money on branded drugs rather than generics. More broadly, employers, health insurers, and/or identity thieves could gain access to users’ profiles, leading to negative consequences, including privacy compromise, social embarrassment, discriminations from employers and insurance companies, identity theft, and so forth [5,20]. Because health-related SNS are not HIPAA-covered entities, these concerns are very real and must be addressed seriously. When users lose their trust and confidence in the ability of a health-related SNS to protect privacy, that company’s reputation will be irreparably damaged.

A Threat Model

For users, a health-related SNS consists of a set of users, a set of mechanisms for exchanging information, a set of binary relations between users, a set of search functions, and a site operator.

The SNS provider and its affiliates may use health information for many purposes. It may also release health information to various third parties and apps or enable exchange of information with other social networks. We may include these additional actors into a usage and sharing network that involves the SNS provider and its affiliates, a set of third parties that collect data from the site, a set of apps that users may invoke within the site, a set of other SNS or websites, and government agencies, including law enforcement and public health. An information flow diagram for a health-related SNS is shown in Figure 1.

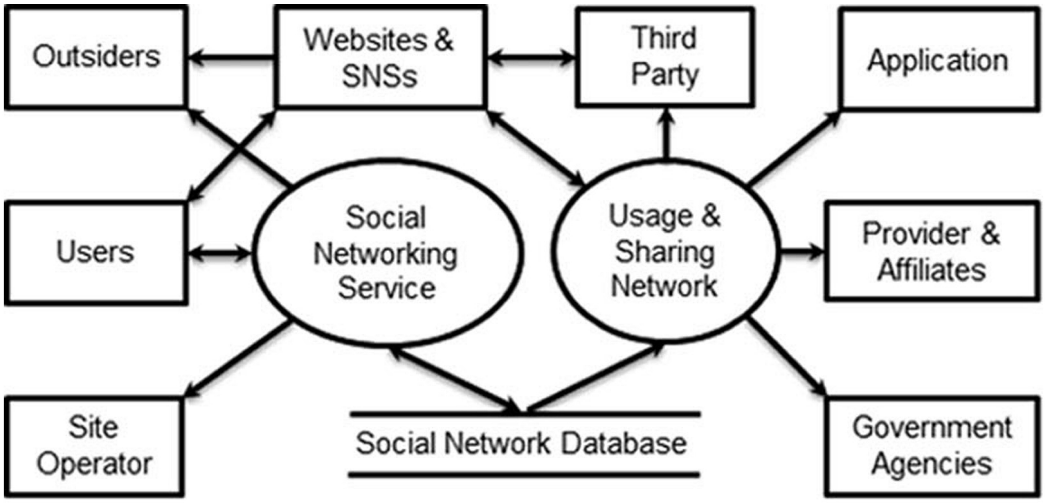
In the social network, a user creates a personal profile, content, and connection network on a health-related SNS. The user may also join other social networks in order to enjoy different social networking services and enter personal health information into other SNS or websites. The site and other websites permanently store the information into their own databases. The site operator uses the information to control the site. The site may make the information visible to other users or even to unintended outsiders including visitors, fake accounts, and attackers. Outsiders may also draw information from other websites.

In the usage and sharing network, the SNS provider and its affiliates may use the accumulated health information for commercial purposes. They may disclose the information to third parties (eg, researchers, marketers, insurers, employers) that may also collect information from other social networks that users have joined and show some information collected

XSL•FO
RenderX

from the current site on other websites. The SNS provider may also permit users to launch various apps that draw information from user profiles in order to create targeted materials. Furthermore, the SNS provider and other SNS providers may share their databases and link different user accounts across multiple SNS due to the collection of more personal health data. Finally, the SNS provider may release the information to government entities for law enforcement or social uses.

Figure 1. A health-related social networking site.



Privacy Threats

Overview

The process of identifying threats to users should recognize users' interest in protecting personal information from parties with which they do not consent or intend to share it. Users are also concerned that personal data may be used in the wrong way or for the wrong purpose. Thus, we look at four key elements in defining a privacy threat—the actors who disclose information, the actors who receive information, the types of information involved, and the purpose. We outline the principal threats of SNS below.

Excessive Revelation of Personal Health Information

Many users have provided unprecedented amounts of detail about their lives, including PII and sensitive health information. Some people hope that exchange of health information will help them access health advice, receive and give social support, manage their conditions, or improve their overall health and quality of life [21,22]. However, health-related SNS may make the information easily accessible to unwanted audiences. Some people may reveal their personal information for the sake of the greater good. Yet they typically have no way of knowing whether their profiles contribute directly to the development of more effective treatments or simply become a lucrative asset for sale. The shared information may contain personal information such as real name and photos, together with their medical conditions. Once personal health information is compromised and the resultant harm is done to that person, it cannot be withdrawn and made private again [20]. Furthermore, users post not just great amounts of private information about themselves but information about other people such as their family members and friends. Although some medical research programs need health information about patients' relatives, disclosing medical information about other people is considered a privacy violation. Individuals sharing information on health trends can, if their submissions are aggregated, reveal information about the health issues affecting their local communities or ethnic groups [4].

Access and Use by Other Users and Visitors

Personal health information may become visible to other users, and visitors may also find the information via the website's search features or even Google searches. This raises the problem of inappropriate access and use by other users and visitors. Even if users can control the access to their own profiles, they may not control what other users and visitors reveal about personal information posted in a public area. For example, other users could be untrustworthy and steal an individual's health information and use it for their own purposes. They may disclose the information to the person's employer or insurer or post it on the Internet.

Secondary Uses and Disclosures by the Social Networking Site Provider and its Affiliates

After users share their personal health information with a health-related SNS, they may lose control over the distribution of their information. The SNS provider has unlimited access to all the information. Ultimately the SNS provider expects that the information will generate insights with considerate scientific as well as economic value. Users are extremely vulnerable because they have little control over the collection, use, and disposition of their information. Privacy can be compromised in many possible ways: targeted advertising, secondary use of the information for research, direct misuse, creation of a permanent record of personal profiles, accidental information release by a site operator, etc.

User Profiling Across Multiple Social Networking Sites

Many users join multiple social networks for different purposes. This means a user may hold multiple profiles, which are stored and shared in different SNS [23-25]. For example, a user creates an account on Facebook mainly to communicate with friends

XSL•FO
RenderX

and families, as well as to share pictures and videos with them. In the meantime, she provides her professional profile and establishes her professional networking on LinkedIn. Furthermore, she stores her personal health records and shares her treatment and symptom on a health-related SNS.

Third parties and companies may use different user accounts and their social relations to connect multiple social networks and produce aggregated user profiles [24,25]. These aggregated user profiles would be immensely valuable to companies looking to market products or services or, in the case of employers, screen potential job applicants. Furthermore, companies can integrate multiple social networks and conduct social network analysis and mining tasks on the integrated social networks [26]. Individual published social network data capture only a partial picture of a user's complete social network. Integration of multiple online and real social networks provides a more complete picture of a user's social network.

Unfortunately, such user profiling and social network integration is not necessarily always beneficial. For example, malicious third parties and identity thieves may use their own crawler systems to obtain a user's private information and friend lists. More seriously, such third parties and individuals could create fake accounts pretending to be this person and then solicit others to connect [25]. These fake accounts can be abused to deliberately leak the user's private information and friend lists to malicious intruders, which could quickly turn into identity theft and fraud, losing a job, hurting relationships, or even worse.

Secondary Uses and Disclosures by Third Parties

The SNS provider may disclose personal health information to third parties and apps. Users cannot assess the risks of divulging personal information unless they know the set of organizations to which their information may be disclosed, and the uses to which it may be put [27]. Because health information is of high commercial value, the accessibility and manipulability of the information creates economic pressures for its use and disclosure for a widening range of commercial and industrial uses. The SNS provider may also allow third-party websites and apps to automatically have access to users' personal information. Data portability technologies may allow many websites and apps to be linked together, letting them share both dynamic content and the nature of the relationships of their users [3,4]. For example, an SNS may communicate with advertising servers, which produce targeted advertising based on details contained in user profiles. The ability to draw data from multiple websites and apps may allow third parties to create a comprehensive digital profile of private data, accumulating more than what a user would have predicted [2,28].

Inability to Detect Sources of Privacy Violations

A health-related SNS cannot assure users' privacy if it lacks automated tracing mechanisms to monitor and track uses and potential misuses of personal information. Visibility and transparency has not been a strong point of health-related SNS. Information mash-ups and the combination of apps and multiple different types of SNS [24] create unexpected information flow through "back channels", impeding users' ability to get a clear view of the way their data are propagating [5]. Different actors (eg, users and apps), linkages, and roles are having dynamic interactions with each other through different ways across multiple social networks or websites. Thus, it is hard for a user to identify the core elements (eg, bridge, hub, broker, power user, proxy) responsible for information dissemination among multiple SNS and find their implicit and explicit relationships with other SNS [24]. Users are often incapable of defending their privacy just because they do not know that their privacy is even endangered. Privacy policies, especially relating to third parties, apps, and social network data sharing and integration, are often vague, uninformative, and seldom reflect users' expectations [2,28,29].

Outsider Attacks

A health-related SNS is vulnerable to attacks from malicious outsiders, such as data scraping and social engineering attacks. Data scraping is a technique that trolls online communities, discussion boards, blogs, and chat rooms looking for personal information that can be used for fraud or any other purposes. For example, data scrapers may choose to work surreptitiously through hidden programs, or they may sign up with a fake email address in order to obtain personal information from unsuspecting users. A patient site also creates a perfect social and ecological environment for spear phishing, viruses, worms, spyware, spoofing, and Web app attacks, facilitated by human vulnerability and easily accessible user profiles [28]. Furthermore, a health-related SNS is vulnerable to social engineering techniques that exploit low entry thresholds to trustful health communities [3].

A Privacy Preservation Model

Overview

Health-related SNS have unique needs to address the principal threats to users and SNS providers not only because personal health information is highly sensitive but also because privacy is essential for building trust, which is the foundational currency of health communications. Today, the dominant approach is a combination of end-user license agreements and privacy settings. Privacy by license agreements is problematic because users have to accept these agreements prior to using SNS services even if they are concerned about privacy. Empirical and theoretical research suggests that users often lack enough information to make privacy-sensitive decisions and, even with sufficient information, are likely to trade off privacy for health benefits [30]. Moreover, the terms of these agreements seldom reflect users' expectations because they can be created and changed only by SNS providers, not by users [29,31].

Current privacy settings provided by most health-related SNS suffer a number of drawbacks. First, since most SNS make "public" their default settings, users may forget to change the default settings. Second, individual self-control is constrained by the user's awareness and education and by the technical design of an SNS, which may impede easy and effective management of settings regarding the access, use, and disclosure of personal information [2]. Furthermore, privacy settings give users control over who sees what on each profile, but they give users little control over what the SNS provider and its affiliates

XSL•FO
RenderX

reveal about them. Therefore, asking individuals to assume full responsibility for policing the use of their profiles by other users and visitors is no longer reasonable, nor does it offer sufficient checks against direct misuse and improper disclosure of personal information by the SNS provider and its affiliates.

Table 2. Privacy threats and countermeasures.

Privacy principles	Privacy threats	Countermeasures
Safe, flexible, and user-friendly privacy settings	Excessive revelation of personal health information; improper access and misuse by other users and visitors; secondary uses and disclosures of personal information; user profiling across multiple SNS.	Banning personally identifiable information; flexible and user-friendly way of setting privacy preferences; individual choice and consent; visualization of connection network; integration of privacy and security settings across multiple SNS.
Privacy by design	Secondary uses and disclosures by the SNS provider and its affiliates; secondary uses and disclosures by third parties; user profiling across multiple SNS.	Sharing de-identified data inside or outside an SNS; limiting use, disclosure, and retention; deleting user accounts upon request; a global privacy preservation model for data sharing and integration across multiple SNS.
Privacy audits	Inability to detect sources of privacy violations; user profiling across multiple SNS.	Audit trails; auditing and monitoring; transparency of data-handling practices; options for users to report privacy invasions; auditing usage and data sharing across multiple SNS.
Security for privacy	Outsider and insider attacks.	Technical barriers such as multifactor authentication, encryption, continuous monitoring, and security analytics; organizational measures such as user education and awareness, options for users to report a security incident, and breach notification and enforcement.

Instead, a privacy model based on a shared responsibility between the SNS provider and users may be better suited as a means of effective protection for both the SNS and its users. User profiles, user-generated content, and social links are the most valuable asset for the SNS provider, and it should be in the best interests of the SNS provider to find solutions to protect those assets through effective means. Therefore, this paper assumes that both the SNS provider and users share the same values concerning protection of user privacy. Direct misuse and improper disclosure of personal information in the usage and sharing network (Figure 1) can lead to conflicting interests for users and the SNS provider. The conflicting interests can be resolved by other means (eg, regulations [2], decentralized social network services, and cryptographic solutions [8-10]) that fall beyond the scope of this paper. The threat analysis outlined above indicates that privacy protection should be considered on four fronts: user self-control, privacy-preserving mechanisms, privacy audits, and security mechanisms. Building on early research [2-7,26,31-33] and the concept of privacy by design [34], this paper proposes a privacy preservation model that incorporates both individual self-protection and privacy-by-design principles. Below we identify key privacy principles and countermeasures to address the principal threats of health-related SNS (Table 2).

Safe, Flexible, and User-Friendly Privacy Settings

Privacy settings play a vital role in matching users' privacy expectations. Many health-related SNS give options to hide certain types of personal information from other users and visitors through the customization of privacy settings. The SNS provider expects users to choose their privacy settings meticulously using available privacy options. But users' self-protection behaviors are constrained by their privacy awareness and by the technical design of privacy settings. Safe, flexible, and user-friendly privacy settings allow the user to set privacy preferences easily and effectively. First, a health-related SNS should turn on privacy settings that limit the collection, display, or sharing of PII by default [3]. For example, the SNS would not make any PII publicly viewable until the user takes affirmative steps to allow this. Second, the SNS can provide flexible privacy settings that afford users fine-grained control over each and every piece of personal information so that other users and visitors cannot access it without explicit consent. Privacy could be compromised by the user's inability to control impressions and manage complex social contexts [7]. It needs to be a major responsibility of the SNS provider to raise the awareness of users and to make its privacy settings very user-friendly. If the SNS enables exchange of information with other SNS or websites, a global model is needed to deal with issues of integration of privacy and security settings across multiple SNS. Third, health-related SNS may provide a means by which users can visualize their current exposure within the community and across multiple social networks. In practice, users have little sense of how their information is accessed and used by other users, visitors, apps, third parties, and other SNS. Graphical displays of the social relations and user accounts linkage across multiple social networks [24,26] would help the user appreciate the potential risks arising from a disclosure and customize their individual settings accordingly.

Privacy by Design

Privacy by design refers to the philosophy and approach of building privacy into the design and architecture of technologies, business practices, and the underlying technical platforms [34]. The presence of protection for users' privacy, including data anonymization and purpose limitation, is crucial to gaining the necessary public trust to make the SNS successful. The following privacy-preserving mechanisms have to be taken into account. First of all, the SNS provider may design architectures that apply appropriate privacy-preserving transformations before transferring the information to individuals and entities. There are several transformation techniques. The safe harbor de-identification method attempts to suppress individual identifiers in order to de-identify the data. Health-related SNS might voluntarily comply with the HIPAA privacy rule by deleting 18 common identifiers before disclosure [35,36]. Under

XSL•FO
RenderX

the HIPAA privacy rule, data are considered de-identified if the covered entity removed the following identifiers from the data: names, addresses, dates, telephone numbers, fax numbers, email addresses, social security numbers, medical record numbers, health plan beneficiary numbers, account numbers, certificate/license numbers, vehicle identifiers and serial numbers (including license plate numbers), device identifiers and serial numbers, Web Universal Resource Locators (URL), Internet Protocol (IP) addresses, biometric identifiers (including finger and voice prints), full-face photographic images and any comparable images, and any other unique identifying number, characteristic, or code.

An alternative approach, known as statistical anonymization techniques [37-39], desensitizes the data by suppressing quasi-identifiers (eg, postal code, birth date, gender, hometown, and/or other demographics), decreasing precision/accuracy, and/or adding confusion to the information in order to make it more difficult to link de-identified data back to the individual. Properly applied statistical anonymization is an effective tool for protecting privacy and preserving the ability to leverage user-generated content for secondary purposes. Furthermore, health-related SNS may use network data anonymization techniques to reduce the identity inference risks from social network data such as social graph, tagging data, email, or instant messaging. The techniques attempt to suppress the user's network structure by graph modification approaches and clustering-based approaches [33]. However, the techniques only allow us to investigate the structural properties of a single anonymized social network. In many cases, node identifiers are essential to link data from different social networks. In order to share useful information among different social networks while protecting privacy, Tang and Yang [26] proposed a generalization and probabilistic approach by generalizing social networks to preserve privacy and integrating the probabilistic models for the generalized social network data for social network analysis and mining.

Over the past few years, however, researchers have found that even de-identified data could be re-identified and attributed to specific individuals [40,41]. Third parties and companies are actively seeking end-user information by linking a variety of different data sources and different user accounts across multiple social networks. The more datasets to which third parties and companies have access, the easier such re-identification becomes. Therefore, the SNS provider and third parties should make a public commitment not to re-identify the data for commercial uses without explicit consent and it should contractually prohibit downstream recipients from doing the same. The SNS may also provide privacy-preserving interfaces for third-party apps while still enabling them to deliver customizable content. Current best practices include "privacy by proxy" mechanisms [32].

Second, the SNS provider may limit the collection, use, and disclosure of personal information to the purposes identified in the privacy notice. Personal information shall not be used or disclosed for purposes other than those identified in the privacy notice, except with consent or where required by law. It is a challenge to find a balance between privacy and utility in data sharing and integration across multiple social networks and websites. On one hand, users' personal information is the most valuable asset for SNS providers and it should also be in their best interest to protect the asset. SNS providers, on the other hand, need to proof their business model by further expanding ways to exploit the value of their users' personal information. Stringent penalties for misuse and improper disclosure of personal information should be established through federal regulations or contractual mechanisms.

Third, a health-related SNS may provide convenient tools to allow users to destroy their profiles and posts completely, in a timely fashion. These tools should allow users to remove their personal information safely and delete or edit their posts in a user-friendly way.

Privacy Audits

Privacy audits provide a means of independently verifying that a health-related SNS operates according to its privacy policies. Auditing and monitoring services are not included in the privacy policies of current health-related SNS. A health-related SNS cannot assure users of their privacy and security unless it enables users to request an "audit trail", detailing when their personal information was accessed, by whom, and for what purpose. A second alternative is to actually audit access and actively notify users in the case of inappropriate access. This principle seeks to assure users that a health-related SNS is operating according to its privacy policies, subject to independent verification. Its component parts and operations are visible and transparent to users. Options for the user to report privacy invasions establish transparency and additional trust in its commitment to adequate treatment of personal information. Furthermore, malicious intruders may use their own crawler systems to obtain a user's private information and friend lists and infer the user account's linkage across multiple SNS and websites. It is highly desirable to design a methodology for auditing usage and data sharing and detecting unauthorized access to each user's personal information across multiple social networks.

Security for Privacy

Health-related SNS may provide appropriate security safeguards that improve privacy. Intruders are increasingly using complicated techniques via the Internet to steal personal information. Traditional security solutions like firewalls and encryption are no longer the centerpiece against social network attacks. Encryption technology for the transmission and storage of personal information provides enhanced security. But data thieves may steal personal information via fake accounts or launch automated crawling and identity theft attacks across different SNS and obtain a large amount of user private information. Some health-related SNS do not use a validation process during new user's registration. Weak authentication of registrants through a functional email address, the preferred validation requirement, is not an adequate method and leads to a proliferation of fake accounts populating the network. Therefore, health-related SNS may develop strong multifactor authentication that combines two or more independent categories of credentials: what the user knows (password), what the user has (security token), and what the user is (a biometric characteristic such as a fingerprint). Health-related SNS may also invest in things like continuous monitoring and security

XSL•FO
RenderX

analytics solutions that monitor the network 24/7 [42], reporting suspicious activities or vulnerability. Social engineering and phishing attacks are the most important threats to users. Sadly, there is no computer program that can protect the network from social engineering or phishing attacks. The best protection is security education and awareness. Health-related SNS could develop proactive communication techniques that raise the level of education and awareness about dangers of privacy and security breaches. Procedures and policies could also be in place for reporting misuse and illegal activity.

The above-mentioned principles and techniques form the basis of how to address the threats of health-related SNS and other eHealth technologies. In principle, many of the techniques and industry best practices needed to implement and enforce these principles are available, if not deployed on existing health-related SNS. We do not have space to detail all the protections for user privacy in this paper, but only to provide a concise set of countermeasures and to relate the countermeasures to the identified privacy threats (Table 2). Since de-identification and informed consent are key elements of privacy laws, these principles and countermeasures can give a health-related SNS legal cover in case of a privacy breach.

Conclusions

A health-related SNS benefits from the increasing amount of personal health information willingly shared on its site, but users are likely to be exposed to privacy and security threats. In this paper, we have developed a threat model that highlights the underlying usage and sharing network behind the SNS and shows the principal threats to users. Because the established solutions like license agreements and unsafe privacy settings are inadequate to mitigate the threats, we proposed a conceptual privacy framework that integrates such foundational principles as safe and flexible setting, privacy by design, privacy audits, and security for privacy. The principles and their associated countermeasures provide a practical way to protect privacy against unauthorized individuals or entities. This proposed model can be generalized to other online settings where personal information is available.

Because personal health information is extremely valuable to both the SNS provider and its business partners, there are always economic pressures on the SNS provider to exploit the value of the database it holds—a prospect that becomes even more tempting if the current business model that supports full user control does not generate sufficient revenue. Hence, there is a tension here, because without effective protections, many users would refrain from sharing health information online due to privacy concerns [43], causing the community to fade away. But if the SNS allows users to keep too much of their information private, there will be less content for creating commercial and social value inside or outside the SNS. Consequently, its business will suffer. The main challenge in the future will be to develop privacy-preserving SNS that protect user privacy while still tapping the richness of user-generated content. All involved parties, and at the foremost the SNS developers, need to understand the potential threats that exist and therefore build privacy and security protections into health-related social networks.

Authors' Contributions

JL contributed to the writing of the paper.

Conflicts of Interest

None declared.

References

1. Angina J, Stecklow S. 'Scrapers' dig deep for data on Web. 2010. Wall Street Journal URL: http://online.wsj.com/article/SB10001424052748703358504575544381288117888.html [accessed 2014-10-25] [WebCite Cache ID 6TarGr4QC]
2. Li J. Privacy policies for health social networking sites. J Am Med Inform Assoc 2013;20(4):704-707 [FREE Full text] [doi: 10.1136/amiajnl-2012-001500] [Medline: 23599228]
3. Li J. Data Protection in Healthcare Social Networks. IEEE Softw 2014 Jan;31(1):46-53. [doi: 10.1109/MS.2013.99]
4. Williams J. Social networking applications in health care: threats to the privacy and security of health information. In: Proceedings of the ICSE Workshop on Software Engineering in Health Care.: ICSE Workshop on Software Engineering in Health Care; 2010 Presented at: ICSE Workshop on Software Engineering in Health Care; 2010; Cape Town, South Africa p. 39-49. [doi: 10.1145/1809085.1809091]
5. Williams J. Weber-Jahnke JH Social networks for health care: Addressing regulatory gaps with privacy-by-design. In: Proceedings of the Eighth Annual International Conference on Privacy Security and Trust. 2010 Presented at: Eighth Annual International Conference on Privacy Security and Trust; 2010; Ottawa, Canada p. 134-143. [doi: 10.1109/PST.2010.5593252]
6. van der Velden M, El EK. "Not all my friends need to know": a qualitative study of teenage patients, privacy, and social media. J Am Med Inform Assoc 2013 Jan 1;20(1):16-24 [FREE Full text] [doi: 10.1136/amiajnl-2012-000949] [Medline: 22771531]
7. Weiss S. Privacy threat model for data portability in social network applications. International Journal of Information Management 2009 Aug;29(4):249-254. [doi: 10.1016/j.ijinfomgt.2009.03.007]
8. Greschbach B, Buchegger S. Friendly surveillance – a new adversary model for privacy in decentralized online social networks. In: Proceedings of the 5th Interdisciplinary Conference on Current Issues in IT Security. 2012 Presented at: 5th Interdisciplinary Conference on Current Issues in IT Security; 2012; Freiburg, Germany p. 5-206.

XSL•FO
RenderX

9. Paul T, Buchegger S, Strufe T. Decentralizing social networking services. In: Salgarelli L, Bianchi G, Blefari-Melazzi N, editors. Trustworthy Internet. Milan: Springer; 2011:187-199.
10. Beato F, Kohlweiss M, Wouters K. Scrambles! Your social network data. In: Proceedings of the 11th International Conference on Privacy Enhancing Technologies, Lecture Notes in Computer Science. Berlin: Springer; 2011:211-225.
11. Inspire Terms and Conditions. URL: https://www.inspire.com/about/terms/ [accessed 2014-10-25] [WebCite Cache ID 6TarQ07dQ]
12. Inspire Principles. URL: https://www.inspire.com/about/our-principles/ [accessed 2014-10-25] [WebCite Cache ID 6TarVPEiE]
13. Inspire Privacy Policy. URL: https://www.inspire.com/about/privacy/ [accessed 2014-10-25] [WebCite Cache ID 6TarZvlpT]
14. CarePages Privacy Policy. URL: https://www.inspire.com/about/privacy/ [accessed 2014-10-25] [WebCite Cache ID 6TarZvlpT]
15. CureTogether Privacy Policy. URL: http://curetogether.com/privacy.php [accessed 2014-10-25] [WebCite Cache ID 6TarhH7HX]
16. DailyStrength Privacy Policy. URL: http://www.dailystrength.org/content/view [accessed 2014-10-25] [WebCite Cache ID 6Tarkte2g]
17. PatientsLikeMe Privacy Policy. URL: http://www.patientslikeme.com/about/privacy [accessed 2014-10-25] [WebCite Cache ID 6Tarp1qSV]
18. Facebook Data Policy. URL: https://www.facebook.com/unsupportedbrowser [accessed 2015-04-29] [WebCite Cache ID 6Y9apWZFy]
19. Twitter Privacy Policy. URL: https://twitter.com/privacy?lang=en [accessed 2015-04-29] [WebCite Cache ID 6Y9auPJtr]
20. Rindfleisch TC. Privacy, information technology, and health care. Commun. ACM 1997;40(8):92-100. [doi: 10.1145/257874.257896]
21. Li J. Improving chronic disease self-management through social networks. Popul Health Manag 2013 Oct;16(5):285-287. [doi: 10.1089/pop.2012.0110] [Medline: 23672233]
22. 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]
23. Jian L, Sahama T. Online multiple profile manager for eHealth information sharing. In: Proceedings of the IEEE International Conference on Communications.: IEEE International Conference on Communications; 2012 Presented at: The 2012 IEEE International Conference on Communications; 2012; Ottawa, Canada p. 3461-3465. [doi: 10.1109/ICC.2012.6364565]
24. Wu-Chen S. Integrating and mining virtual communities across multiple online social networks: Concepts, approaches and challenges. In: Proceedings of the Fourth International Conference on Digital Information and Communication Technology and its Applications. 2014 Presented at: Fourth International Conference on Digital Information and Communication Technology and its Applications; 2014; Bangkok, Thailand p. 199-204. [doi: 10.1109/DICTAP.2014.6821682]
25. Shen Y, Jin H. Controlled information sharing for user accounts linkage across multiple online social networks. In: Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management. 2014 Presented at: 23rd ACM International Conference on Conference on Information and Knowledge Management; 2014; Shanghai, China p. 23-390. [doi: 10.1145/2661829.2661985]
26. Tang X, Yang CC. Social network integration and analysis using a generalization and probabilistic approach for privacy preservation. Security Informatics 2012;1(1):7. [doi: 10.1186/2190-8532-1-7]
27. Krishnamurthy B, Wills C. Characterizing privacy in online social networks. In: Proceedings of the First Workshop on Online Social Networks. 2008 Presented at: First Workshop on Online Social Networks; 2008; Seattle p. 37-42. [doi: 10.1145/1397735.1397744]
28. ENISA. 2007. Security issues and recommendations for online social networks URL: https://www.enisa.europa.eu/publications/archive/security-issues-and-recommendations-for-online-social-networks [accessed 2014-10-25] [WebCite Cache ID 6Tart80yg]
29. McGraw D, Dempsey JX, Harris L, Goldman J. Privacy as an enabler, not an impediment: building trust into health information exchange. Health Aff (Millwood) 2009;28(2):416-427 [FREE Full text] [doi: 10.1377/hlthaff.28.2.416] [Medline: 19275998]
30. Acquisti A, Grossklags J. Privacy and rationality in individual decision making. IEEE Secur. Privacy Mag 2005 Jan;3(1):26-33. [doi: 10.1109/MSP.2005.22]
31. Li J. Ensuring Privacy in a Personal Health Record System. Computer 2015 Feb;48(2):24-31. [doi: 10.1109/MC.2015.43]
32. Felt A, Evans D. Privacy protection for social networking APIs. In: Proceedings of the 2008 IEEE Symposium on Security and Privacy. 2008 May 18-21 Presented at: IEEE Symposium on Security and Privacy; 2008; Oakland, CA p. 2008-2008.
33. Zhou B, Pei J, Luk W. A brief survey on anonymization techniques for privacy preserving publishing of social network data. SIGKDD Explor. Newsl 2008 Dec 20;10(2):12. [doi: 10.1145/1540276.1540279]
34. Cavoukian A. 7 Foundational Principles - Privacy by design. URL: http://www.privacybydesign.ca/index.php/about-pbd/7-foundamental-principles [accessed 2014-10-25] [WebCite Cache ID 6TarwriQC]
35. McGraw D. Building public trust in uses of Health Insurance Portability and Accountability Act de-identified data. J Am Med Inform Assoc 2013 Jan 1;20(1):29-34 [FREE Full text] [doi: 10.1136/amiajnl-2012-000936] [Medline: 22735615]

XSL•FO
RenderX

36. El EK, Arbuckle L, Koru G, Eze B, Gaudette L, Neri E, et al. De-identification methods for open health data: the case of the Heritage Health Prize claims dataset. J Med Internet Res 2012 Feb;14(1):e33 [FREE Full text] [doi: 10.2196/jmir.2001] [Medline: 22370452]
37. Malin B, Benitez K, Masys D. Never too old for anonymity: a statistical standard for demographic data sharing via the HIPAA Privacy Rule. J Am Med Inform Assoc 2011;18(1):3-10 [FREE Full text] [doi: 10.1136/jamia.2010.004622] [Medline: 21169618]
38. El EK, Dankar FK, Issa R, Jonker E, Amyot D, Cogo E, et al. A globally optimal k-anonymity method for the de-identification of health data. J Am Med Inform Assoc 2009;16(5):670-682 [FREE Full text] [doi: 10.1197/jamia.M3144] [Medline: 19567795]
39. Mohammed N, Jiang X, Chen R, Fung BC, Ohno-Machado L. Privacy-preserving heterogeneous health data sharing. J Am Med Inform Assoc 2013 May 1;20(3):462-469 [FREE Full text] [doi: 10.1136/amiajnl-2012-001027] [Medline: 23242630]
40. Bobicev V, Sokolova M, El EK, Jafer Y, Dewar B, Jonker E, et al. Can anonymous posters on medical forums be reidentified? J Med Internet Res 2013;15(10):e215 [FREE Full text] [doi: 10.2196/jmir.2514] [Medline: 24091380]
41. Narayanan A, Shmatikov V. Robust De-anonymization of Large Sparse Datasets. In: Proceedings of the 2008 IEEE Symposium on Security and Privacy. 2008 Presented at: 2008 IEEE Symposium on Security and Privacy; 2008; Oakland, CA p. 111-125.
42. SANS Institute. A real-time approach for continuous monitoring URL: http://www.sans.org/reading-room/whitepapers/analyst/real-time-approach-continuous-monitoring-34950 [accessed 2014-10-25] [WebCite Cache ID 6Tas04tEW]
43. 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]

Abbreviations

HIPAA: Health Insurance Portability and Accountability Act
HITECH: Health Information Technology for Economic and Clinical Health Act
IP: Internet protocol
PII: personally identifiable information
SNS: social networking sites
URL: uniform resource locator

Edited by G Eysenbach; submitted 25.10.14; peer-reviewed by WC Su, M van der Velden; comments to author 19.02.15; revised version received 03.05.15; accepted 24.05.15; published 08.07.15.

Please cite as:
Li J
A Privacy Preservation Model for Health-Related Social Networking Sites
J Med Internet Res 2015;17(7):e168
URL: http://www.jmir.org/2015/7/e168/
doi:10.2196/jmir.3973
PMID:26155953

©Jingquan Li. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 08.07.2015. 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

Epidemic Wave Dynamics Attributable to Urban Community Structure: A Theoretical Characterization of Disease Transmission in a Large Network

Anne G Hoen[1,2,3*], PhD; Thomas J Hladish[4,5*], PhD; Rosalind M Eggo[5], PhD; Michael Lenczner[6], BSc; John S Brownstein[1,2*], PhD; Lauren Ancel Meyers[5,7*], PhD

[1]Computational Epidemiology Group, Children's Hospital Informatics Program, Boston Children's Hospital, Boston, MA, United States

[2]Department of Pediatrics, Harvard Medical School, Harvard University, Boston, MA, United States

[3]Department of Epidemiology, Geisel School of Medicine at Dartmouth, Dartmouth College, Hanover, NH, United States

[4]Emerging Pathogens Institute, Department of Biology, University of Florida, Gainesville, FL, United States

[5]Department of Integrative Biology, The University of Texas at Austin, Austin, TX, United States

[6]Île Sans Fil, Montreal, QC, Canada

[7]Santa Fe Institute, Santa Fe, NM, United States

[*]these authors contributed equally

Corresponding Author:
Thomas J Hladish, PhD
Emerging Pathogens Institute
Department of Biology
University of Florida
PO Box 118525
Gainesville, FL, 32611
United States
Phone: 1 512 481 2403
Fax: 1 352 392 3704
Email: tjhladish@ufl.edu

Abstract

Background: Multiple waves of transmission during infectious disease epidemics represent a major public health challenge, but the ecological and behavioral drivers of epidemic resurgence are poorly understood. In theory, community structure—aggregation into highly intraconnected and loosely interconnected social groups—within human populations may lead to punctuated outbreaks as diseases progress from one community to the next. However, this explanation has been largely overlooked in favor of temporal shifts in environmental conditions and human behavior and because of the difficulties associated with estimating large-scale contact patterns.

Objective: The aim was to characterize naturally arising patterns of human contact that are capable of producing simulated epidemics with multiple wave structures.

Methods: We used an extensive dataset of proximal physical contacts between users of a public Wi-Fi Internet system to evaluate the epidemiological implications of an empirical urban contact network. We characterized the modularity (community structure) of the network and then estimated epidemic dynamics under a percolation-based model of infectious disease spread on the network. We classified simulated epidemics as multiwave using a novel metric and we identified network structures that were critical to the network's ability to produce multiwave epidemics.

Results: We identified robust community structure in a large, empirical urban contact network from which multiwave epidemics may emerge naturally. This pattern was fueled by a special kind of insularity in which locally popular individuals were not the ones forging contacts with more distant social groups.

Conclusions: Our results suggest that ordinary contact patterns can produce multiwave epidemics at the scale of a single urban area without the temporal shifts that are usually assumed to be responsible. Understanding the role of community structure in epidemic dynamics allows officials to anticipate epidemic resurgence without having to forecast future changes in hosts, pathogens, or the environment.

XSL•FO
RenderX

(J Med Internet Res 2015;17(7):e169) doi:10.2196/jmir.3720

KEYWORDS

communicable diseases; epidemics; transmission

Introduction

Epidemics of infectious diseases are frequently characterized by multiple waves of infection [1-3]. Notably, the 1918 influenza pandemic spread through several US and European cities in multiple waves with local variation in the frequency and timing of individual epidemic peaks [4-8]. Predicting when and where disease will resurge is critical to effective prevention and control. However, the drivers and dynamics of multiwave epidemics are unclear. For influenza pandemics, possible explanations include antigenic drift [8-12], waning immunity [13], changing environmental conditions [12,14,15], and social distancing behavior [15-17].

Community structure—aggregation into highly intraconnected but loosely interconnected groups—is a common feature of social contact networks [18] that can potentially drive multiwave epidemics as a disease spreads through one group before emerging in another. However, community structure has been neglected as a possible explanation for multiwave influenza pandemics, in part because it is difficult to detect and estimate [19]. Most studies describing routine human contact patterns have relied on diary- or questionnaire-based surveys [20] or specially deployed wireless sensors [21] and, thus, rarely yield data sufficient for inferring large-scale aggregations. Social networks estimated from electronic "contacts" (ie, cell phones, social networking websites) have been shown to exhibit community structure at larger scales [22-26], but do not capture the physical interactions through which diseases spread. However, the ubiquity of community structure across these networks suggests that it may be a general hallmark of social networks.

Here, we address the hypothesis that contact patterns in a large, empirical, urban contact network are sufficient to generate multiwave epidemics for pandemic influenza-like diseases in the absence of any temporal changes in the hosts, pathogen, or environment. We find that the fate of an epidemic in such a network—whether and when multiple waves occur—depends not only on community structure but also, critically, the presence or absence of bridge superspreaders who forge connections between communities. Direct links between the popular members of different communities synchronize outbreaks; the occasional absence of such bridges provides the epidemiological separation underlying multiwave epidemics.

Interactions between strangers can serve as critical transmission routes for respiratory diseases such as influenza, yet they are difficult to capture in traditional sociological surveys. Using data indicating the physical proximity of more than 100,000 Wi-Fi hotspots users, we characterize the structure of an urban extrasocial interaction network and assess its epidemiological implications.

Methods

Data

Île Sans Fil (ÎSF) is a not-for-profit organization established in 2004 in Montreal, Canada, that operates a system of public Internet hotspots. Hotspots are located in cafes, community and recreation centers, salons, markets, and other small businesses and public places. They are maintained by ÎSF staff and volunteers with the Internet connection provided by the establishment. We analyzed the database of all connections to the system of 352 hotspots between August 2004 and March 2010. Raw data from the ÎSF database consisted of 2.18 million connection records. Each record included an anonymized user ID, latitude and longitude coordinates for each ÎSF hotspot location, connection and disconnection times, and the unique media access control address for the user's wireless device. The data reported in this paper are available from the Community Resource for Archiving Wireless Data at Dartmouth (CRAWDAD) archive [27].

Network Construction

We built a contact network by interpreting each individual user as a node and concurrent ÎSF usage at the same hotspot as an edge. This preliminary network contained 114,810 nodes and 1.2 million edges. It contained both self-loops (users connecting multiple devices at once) and parallel edges (pairs of users with multiple overlapping hotspot visits) that we removed to produce a nonredundant network with 637,430 edges. We analyzed the largest connected component of this network, which consisted of 103,425 nodes and 630,893 edges.

Community Structure Analysis

Modularity (Q) quantifies the extent of community structure in a network relative to a comparable random network. Given a network and a particular partitioning of the nodes into communities, Q is defined as the number of edges contained within communities minus the number of edges expected to fall within communities if the edges were distributed randomly (preserving the degrees of all nodes), normalized for network size. Q ranges from zero for randomly connected networks to greater than 0.3 for networks with substantial community structure [28]. We used a heuristic method to divide the Montreal network into a set of communities that maximized Q using an algorithm [28] that initially assigned each node to its own community and then iteratively aggregated whichever pair of communities resulted in the largest increase in Q. We identified 1420 distinct communities associated with a Q value of 0.69.

Epidemic Simulations

Epidemic curves were simulated by EpiFire [29] using a chain-binomial [30] network epidemic simulator with a susceptible-infectious-recovered state progression. Each epidemic begins with all nodes susceptible except for a single,

XSL•FO
RenderX

randomly selected, infectious node. Nodes remain infectious for a fixed period with arbitrary time units, after which they become recovered for the remainder of the simulation. Transmission from infectious nodes to susceptible neighbors is attempted once per time unit with transmission probability T_{cb}. We assumed an infectious period of 5 time units because it yielded sufficient temporal resolution to distinguish epidemic waves. T_{cb} relates to the percolation transmission probability as given by $T=1-(1-T_{cb})^d$, where *d* is the infectious period. We simulated epidemics across 3 transmission scenarios defined by the value of the basic reproduction number, R_0, which is defined as the expected number of secondary cases produced by a typical infection in a completely susceptible population: a low R_0=1.9, similar to estimates for recent influenza pandemics [31-33]; a moderate R_0=2.4, which maximizes the probability of multiwave epidemics and is in the range estimated for the 1918 influenza pandemic in the United States [34,35]; as well as a high R_0=7.5, where spread is rapid both within and between communities. Finally, we performed a sensitivity analysis with respect to the length of the infectious period and considered values of 1, 5, 10, and 50 time steps.

Detecting Multiwave Epidemics

To automatically identify epidemic curves exhibiting multiple waves, we defined a new 2-peak metric (*TP*), indicating the depth of the deepest valley in the epidemic time series (specifically, the geometric mean of the heights of on each side of the deepest valley). Additional details are provided in Multimedia Appendix 1. Example epidemic curves and their corresponding value of *TP* are shown in the supplementary information (see Figure S1 in Multimedia Appendix 2). The distribution of *TP* values tended to be bimodal with single-wave epidemics yielding values close to zero and multiwave epidemics yielding higher values (see Figure S2 in Multimedia Appendix 2).

Percolation-Based Approximations of R_0, Epidemic Size, and Community Bridging

We adapted methods from percolation theory [36] to estimate global and community-specific values of R_0, final epidemic sizes, and epidemiological connectivity among different communities to test the hypothesis that multiwave epidemics occur in the absence of between-community degree assortativity. The details and derivations of these percolation-based quantities can be found in Multimedia Appendix 1.

Network Shuffling

To examine the epidemiological impact of community structure we constructed "null" networks that shared many properties of the original Montreal network (eg, the degree distribution) but lacked community structure. Specifically, we iteratively randomized the network by shuffling connections; that is, we chose random pairs of connections and swapped the ends (eg, A-B, C-D became A-D, C-B). This slowly degraded community structure while preserving the number of contacts for each individual. We selected a fraction *f* of edges and broke them to form stubs (half-edges still attached to their nodes). Then, the list of stubs was randomized and each sequential pair of stubs was connected. We eliminated newly formed self-edges and redundant edges via edge swaps with randomly chosen edges, leading to randomization of slightly more than the intended fraction of edges (~1.01*f*). Modeling each community as a semirandom network with the observed within- and between-community degree distributions, we estimated the impact of random shuffling and variation in R_0 on the epidemiological proximity of communities. We built 2048 new networks by randomly shuffling from 0.1% of edges (631 edges) to 3% of edges (18,927 edges) in increments of 0.1%. All estimates are averages based on the networks built for each shuffling level.

Results

We analyzed the network of more than 600,000 physically proximal contacts between 103,425 users of a free public Wi-Fi hotspot system in Montreal, Canada (hereafter the Montreal network) to examine the effects of ordinary urban contact patterns on epidemic wave dynamics. We used an established heuristic method [28] to divide the Montreal network into 1420 distinct communities (Figure 1; also see Figure S3 in Multimedia Appendix 2). The 3 largest communities together contained 82,228 of 103,425 (79.50%) users in the network, with 38,569 (community I), 28,101 (community II), and 15,558 (community III) users. The mean degree (number of contacts) per user in each of these communities was 13.4 (SD 40.2), 8.3 (SD 23.7), and 26.7 (SD 76.4), respectively, compared with a mean of 4.6 (SD 9.3) for the 21,197 (20.50%) remaining users. The communities exhibited distinct geographic signatures corresponding to large mixed commercial and residential areas in the city, with considerable overlap occurring in Downtown Montreal (Figure 2).

R_0 is related to the likelihood and extent of a sustained outbreak [37]. R_0 depends on the transmissibility of the pathogen, host recovery, and the structure of the host contact network [38]. Assuming that within-community contacts are approximately random and using a percolation-based model [39], we estimated that a disease with a global R_0 equal to 1 had local R_0 values of 0.8, 0.4, and 1.6 in communities I, II, and III, respectively (when considering only within-community edges) and exhibited considerable variability in epidemiological vulnerability across communities (Figure 3).

We simulated epidemics through the Montreal network with a stochastic susceptible-infected-recovered model [37] across a range of R_0 values. At low R_0, only 2 of 3 communities (I and III) sustained transmission, whereas at high R_0, epidemic spread was relatively synchronized between communities. Under both of these scenarios, multiwave epidemics were possible but relatively infrequent (Figure 3). At an intermediate value, in the range estimated for the 1918 influenza pandemic in the United States (R_0=2.4) [34,35], 44.60% (446/1000) of all epidemics exhibited multiple waves and 87.9% (392/446) of these had an initial epidemic wave in community III with a subsequent wave dominated by the 2 larger communities (Figures 4 and 5). When the first wave was dominated by community III, its peak occurred a mean of 27 time steps (SD

XSL•FO
RenderX

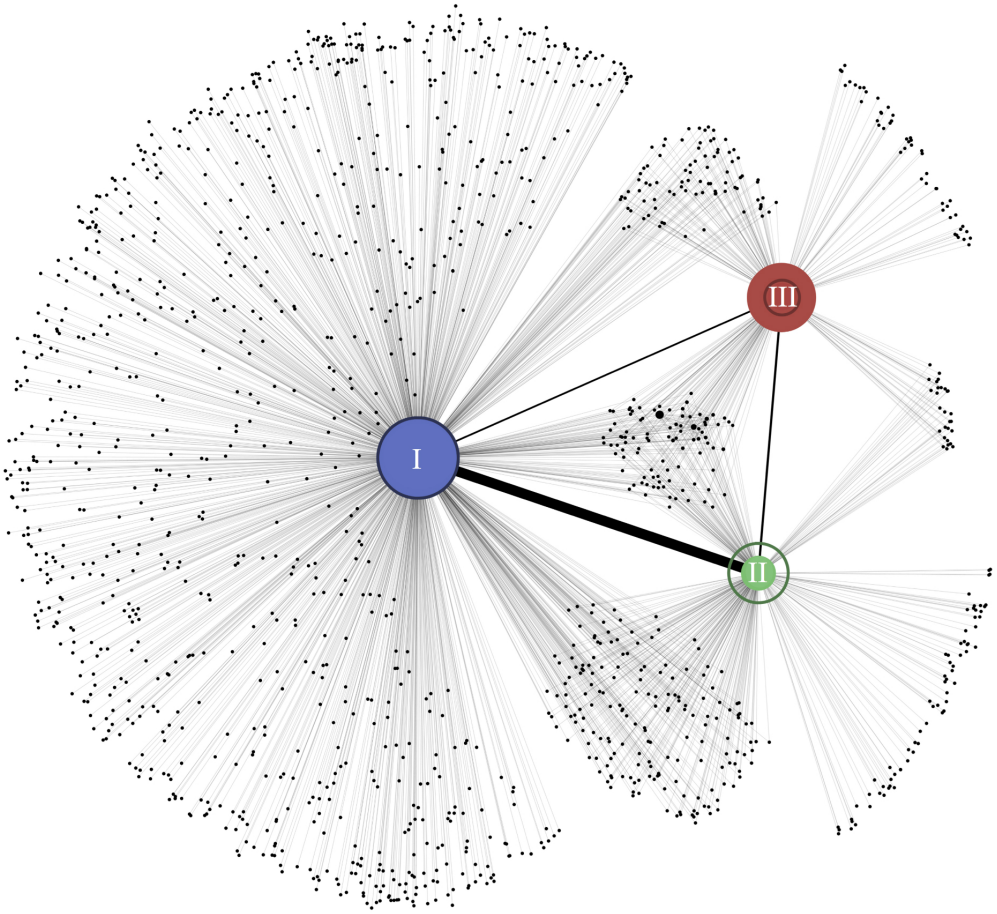
7; n=4074) before that of the second wave. When the second wave was dominated by community III, its peak lagged behind the peak of the first wave by a mean of 22 time steps (SD 8; n=572). The relative size of the second wave increased with R_0 because community III became epidemiologically saturated more quickly than the other 2 communities (Figures 3 and 4). Sensitivity analysis suggested that these results are robust to the length of the infectious period (see Figure S4 in Multimedia Appendix 2).

A modest amount of network shuffling (<3% of edges) almost entirely eliminated multiwave epidemics (Figure 6), whereas it minimally impacted the extent of community structure according to standard metrics including Q (Figure 7) and the numbers of edges linking distinct communities (see Figure S5 in Multimedia Appendix 2). This suggests that the critical driver of 2-peaked epidemics is not the number but the nature of intermodule contacts. We hypothesized that epidemiological synchrony arises when locally popular users in one community tended to be connected to locally popular users in another (between-community degree assortativity) and multiwave epidemics can occur only when communities lack such connectivity.

We formalized and tested this idea by assuming, again, that within-community edges form semirandom networks and used new percolation-based estimates to characterize the epidemiological bridges between the 3 major communities. Given the degree distribution of the Montreal network (see Figure S6 in Multimedia Appendix 2), we found that the number of users in community III expected to form epidemiological bridges to community I increased rapidly with shuffling, whereas communities I and II were tightly connected by bridging individuals in the original network and this connection persisted through shuffling (Figure 8). Shuffling also led to a rapid decrease in the probability that an epidemic starting in community III would spark an epidemic in community I sufficiently late to appear 2-peaked (Figure 9). The precipitous decline in two-wave epidemics with shuffling coincides with the rapid creation of epidemiological bridges and decrease in the expected waiting time between community outbreaks.

The rapid deisolation of the internally well-connected community III occurred because random shuffling targeted users proportional to their degree, which tended to connect high-degree individuals inside community III with high-degree users from elsewhere in the full network. The locally popular but highly insular users of community III thereby quickly formed bridges to popular users in the other major communities.

Figure 1. Visualization of connectivity between 1420 communities identified at maximum modularity (Q). Each circle represents a community with filled circle diameter indicating the relative number of within-module edges. Lines joining pairs of communities are drawn with a thickness that is proportional to the number of edges connecting them. The largest 3 communities (community I, II, and III) are labeled and filled in color. Darker rings superimposed on communities I, II, and III are proportional in diameter to the number of nodes making up each.



XSL•FO
RenderX

Figure 2. Map of hotspot locations with points on map represented by pie charts indicating the relative contributions of each community to the total visits recorded at that hotspot. Community III was primarily concentrated in the Gay Village neighborhood of the Ville-Marie borough of Montreal, whereas communities I and II primarily occupied the high-traffic commercial areas on either side of the Plateau-Mont-Royal neighborhood; all 3 communities coincided downtown. Each grid square is colored to represent the locally dominant community. Squares with no hotspots are colored to represent the dominant community at the nearest hotspot.

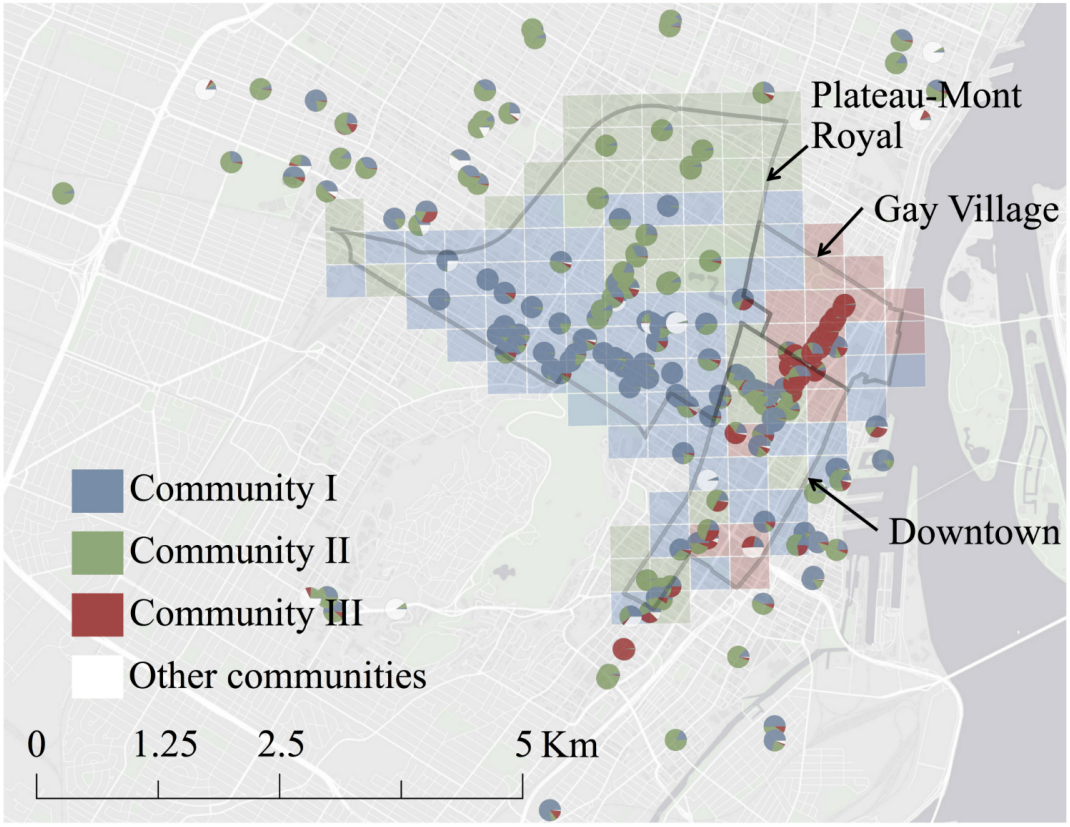
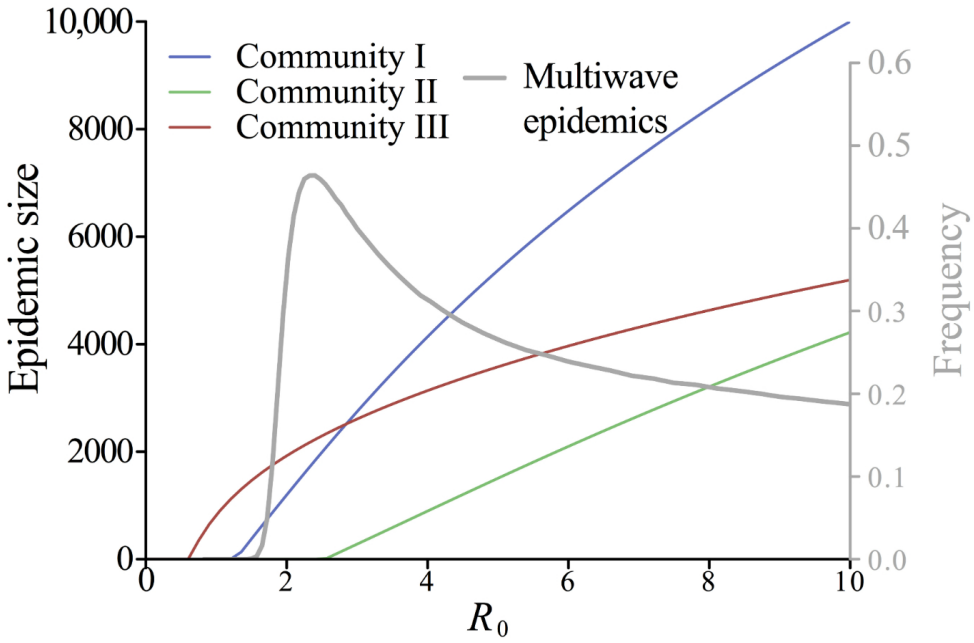


Figure 3. Expected within-community epidemic size assuming that communities were approximately random networks and maintained their empirical within-community degree distributions (colored lines; primary y-axis). The epidemic threshold for each community (ie, R_0 value for which transmission is sustained) is lowest for community III, followed by communities I and II. The frequency of multiwave epidemics depended on R_0 and is highest when R_0=2.4 (gray line; secondary y-axis).



XSL•FO
RenderX

Figure 4. Taxonomy of epidemic curves. For each R_0, 1000 simulated epidemics were classified as either single-wave (left), multiwave starting with a community III wave followed by a community I and II wave (middle), or multiwave ending in a community III wave (right). At R_0=2.4, 554 of 1000 (55.40%) had a single wave and 446 of 1000 (44.60%) had 2 waves; of the 446 with 2 waves, community III dominated the first wave in 392 (87.89%). At R_0=1.9 and R_0=7.5, only 22.20% (222/1000) and 20.60% (206/1000) of epidemics exhibited 2 waves, respectively. Time series are superimposed so that the peaks of the largest waves align.

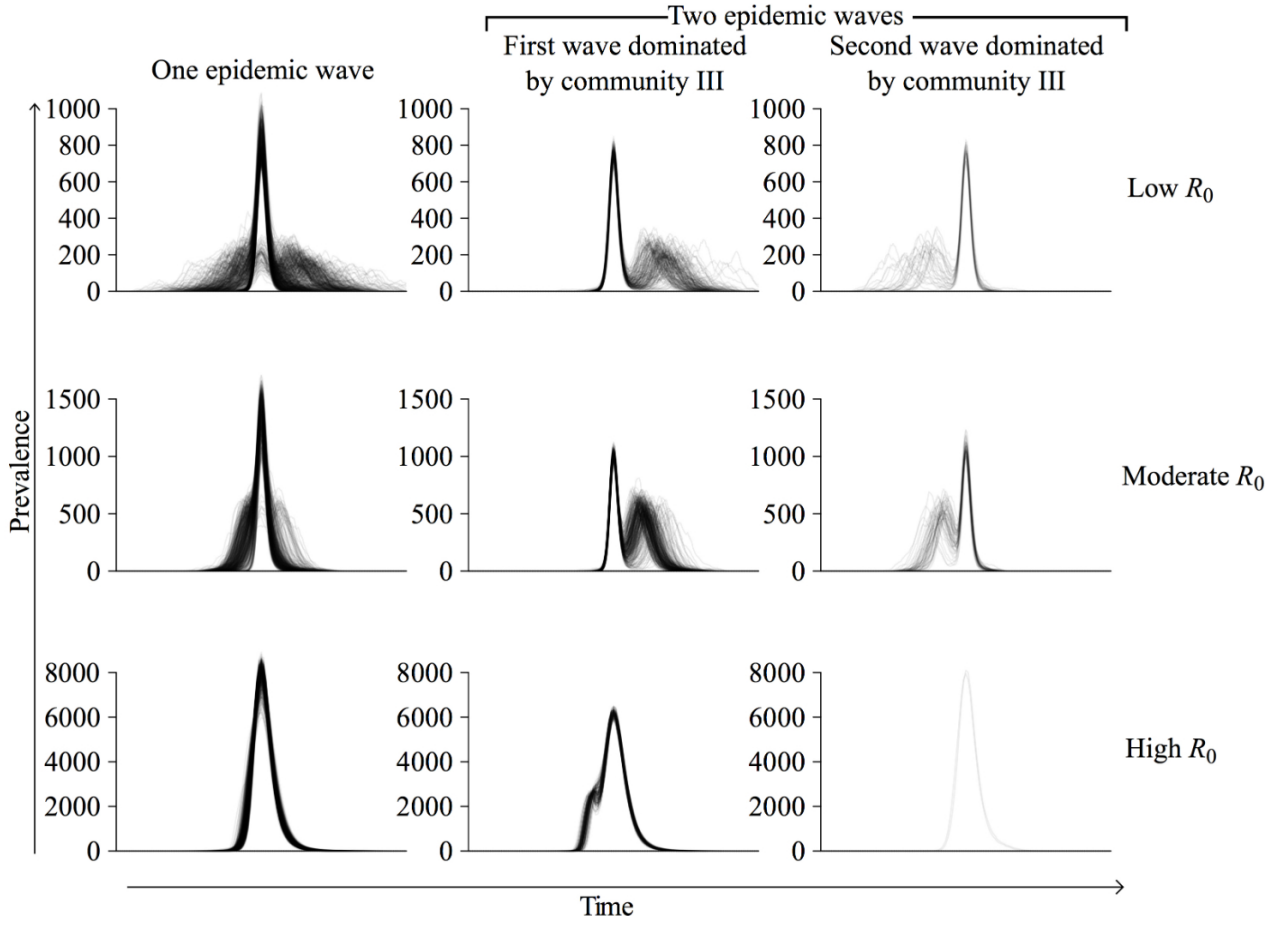
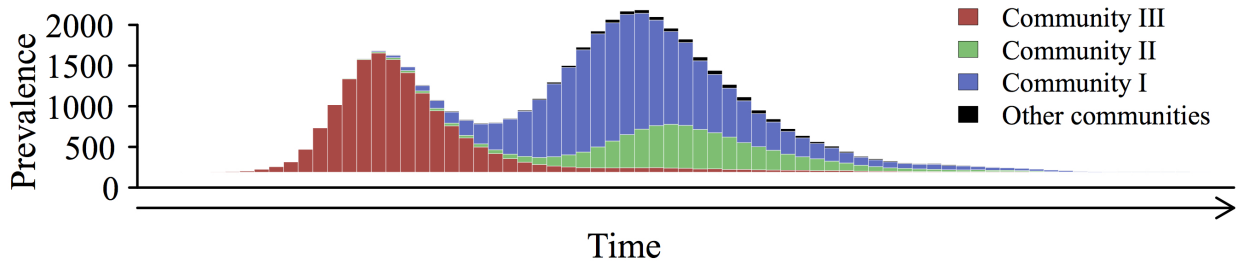


Figure 5. A typical epidemic curve with 2 waves (R_0=3.7). Community III drove the first wave; communities I and II drove the second wave.



XSL•FO
RenderX

Figure 6. The frequency of multiwave epidemics at varying values of R_0 and network shuffling with warm colors indicating a higher proportion of multiwave epidemics and cool colors indicating a low proportion of multiwave epidemics (frequency values indicated on contours). Frequencies were calculated across all epidemics (top left) and stratified by starting community (top right and bottom). Each pixel is based on 81,920 simulated epidemics originating in the specified community (10 simulations on each of 8192 uniquely shuffled networks).

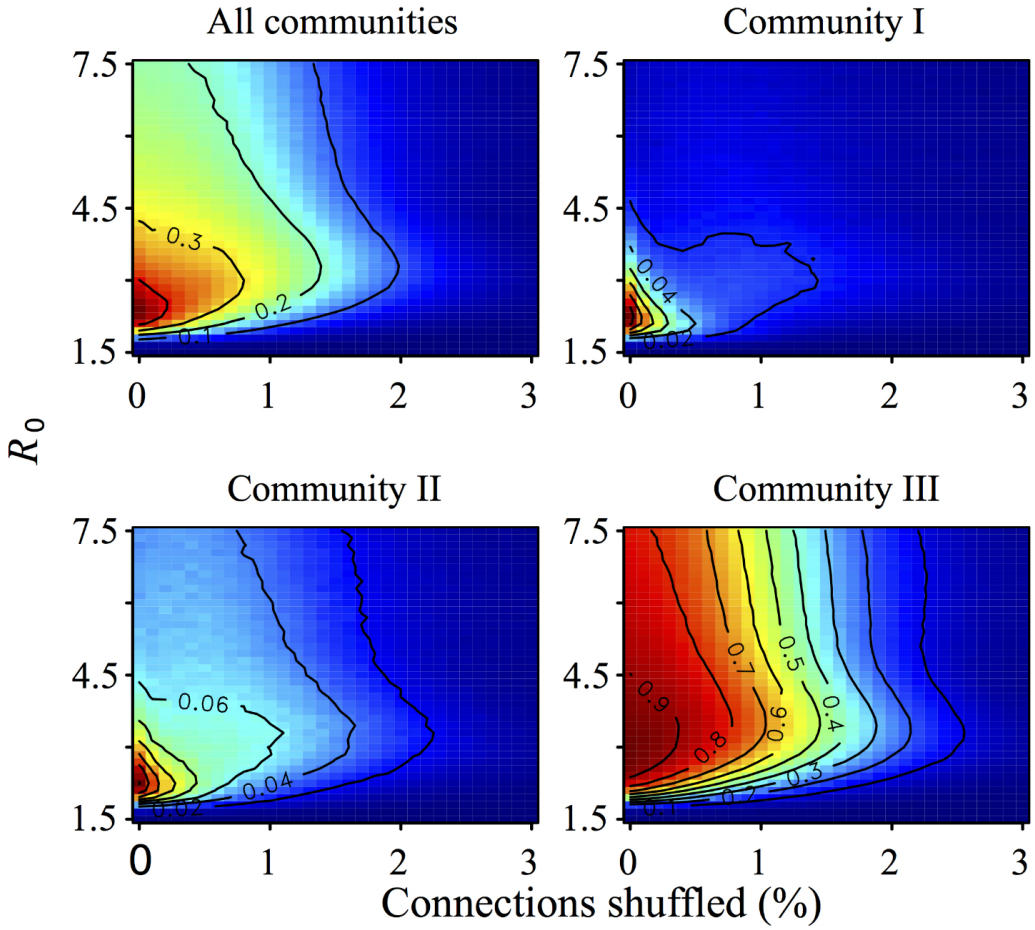
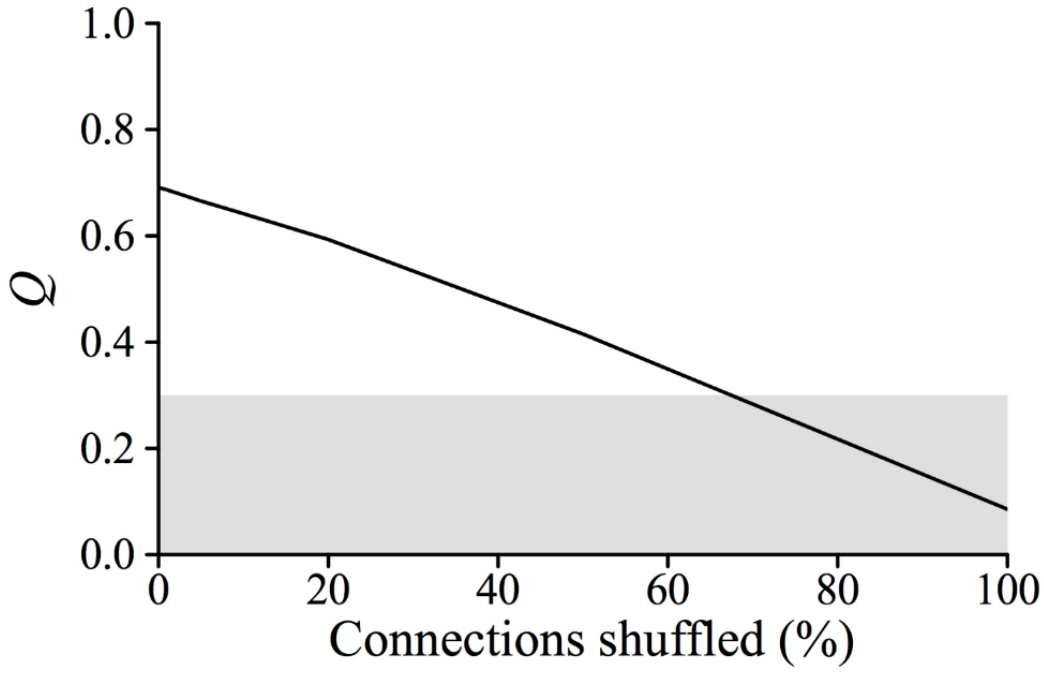


Figure 7. Relationship between modularity (Q) and network shuffling. Shaded area indicates the conventional community structure threshold of 0.3. Note larger x-axis range in this figure compared with other figures.



XSL•FO
RenderX

Figure 8. The proportion of nodes in community II (gray line) and community III (black line) with the ability to spark an epidemic (i.e., φ >0; see Multimedia Appendix 1) in module I across varying levels of network shuffling at R_0=2.4.

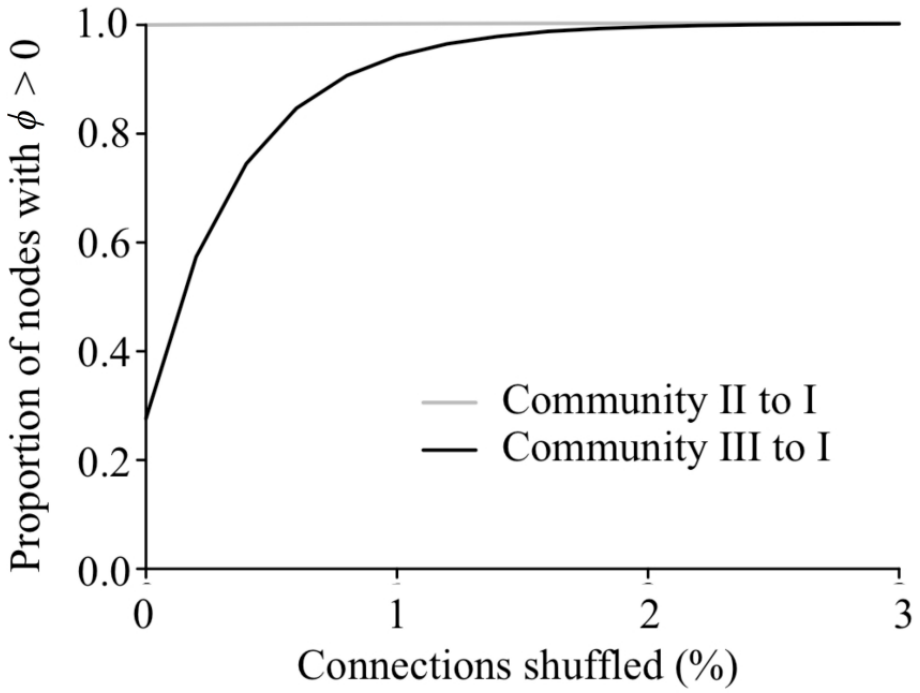
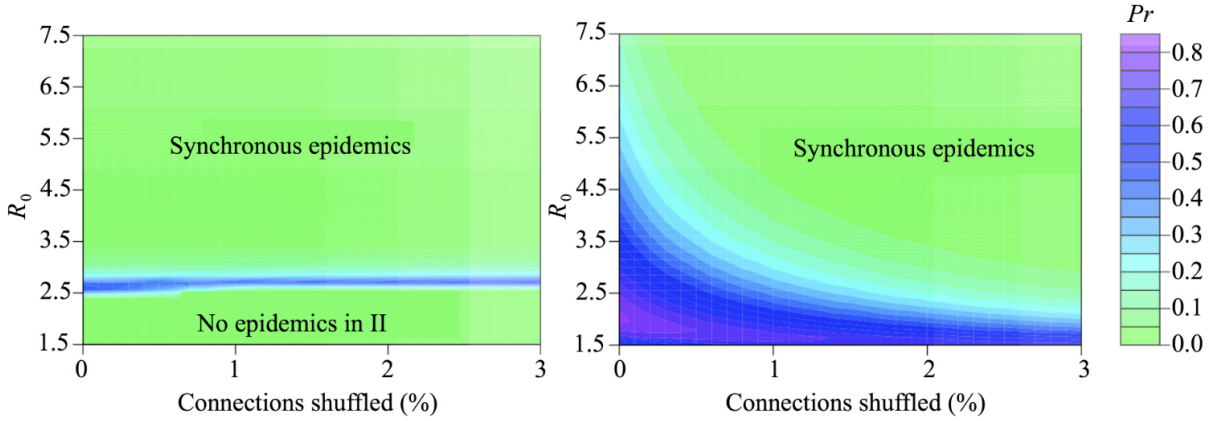


Figure 9. The probability that one community will seed an epidemic in another sufficiently late to produce an asynchronous secondary wave is shown for epidemics spreading from community I to II (left) and community III to I (right).



Discussion

Principal Findings

Community structure is a prominent feature of the Montreal network, with the 3 major communities exhibiting substantially different epidemiological thresholds and dynamics. However, community structure alone is insufficient to produce multiwave epidemics because direct contacts between highly connected individuals in different communities can fuel rapid intercommunity transmission. It is the insularity of the most intraconnected community that makes multiple waves not just possible, but likely, when R_0 resembles that estimated for the 1918 influenza pandemic. The Montreal network shows anecdotally that both strongly connected and insular communities can emerge naturally in human social systems and exist side-by-side. However, insularity is fragile, disappearing with minimal perturbation to the network. In contrast, the presence of strong bridges between communities is more robust and arises by random (or other) processes that link highly connected individuals to one another. Conversely, we expect that social differences, geographic heterogeneity in contact patterns (eg, city vs suburbs), and polarizing events would generate insularity.

The 3 large Montreal hotspot communities exhibit substantially different epidemiological characteristics, including epidemic thresholds and sensitivity to the reproduction number. For mildly transmissible diseases just above the epidemic threshold, some communities may be fully protected by the sparseness of their within-community network. Well above the epidemic threshold, the attack rate may vary considerably among communities, depending again on the structure of their within-community network.

XSL•FO
RenderX

Ball and Neal [40] introduced a theoretical framework addressing the epidemiological interplay between local and global connectivity, and showed that transient long-range connections can fuel epidemics even when local network structure is too sparse to sustain epidemics. Although their model considers transient global contacts rather than fixed community structure, a similar approach may ultimately provide a theoretical perspective on the epidemiological phenomena we observed in the Montreal network.

The insularity of the most intraconnected Montreal community is likely to produce a multiwave epidemic when an outbreak originates in that community and the R_0 is moderate, close to that estimated for the 1918 influenza pandemic. For less contagious diseases, epidemics are unlikely to escape beyond this community; for highly contagious diseases, epidemics flow readily among communities. Our results thus support an alternative nontemporal explanation for multiwave epidemics—insular community structure—that arises naturally from social network structure fundamental to human interactions.

All historical influenza pandemics have produced multiple epidemic waves in many North American cities [1]. Specifically, Montreal experienced 2 distinct waves of influenza-related mortality during the 1957-1958 pandemic, which did not occur elsewhere in Quebec. McDonald [41] attributed this difference to population density. The 2009 pandemic also produced dual waves in Montreal [42]. Prior attempts to explain these patterns have largely overlooked population structure in favor of dynamic phenomena, such as pathogen evolution or host behavior [14,15,43,44]. This omission stems partly from the chronic lack of data and insight into the structure of urban contact networks. Although this remains a challenge, our study suggests that urban community structure may naturally produce multiwave pandemics, even without any temporal forcing.

Limitations

The major limitation of our study lies in the inherent difficulty of capturing and characterizing human contact patterns at an individual level. The Montreal ÎSF network excludes many interactions important to disease transmission, such as those occurring in homes and schools. In addition, some of the colocation contacts in our dataset may not have been sufficient for disease transmission. For these 2 reasons, the ÎSF network does not fully reflect the human contact patterns responsible for disease spread; however, it reveals previously uncharacterized urban-scale community structure that may reflect fundamental geographic, economic, and cultural processes that shape physical contact networks.

Conclusions

Anticipating the emergence of secondary epidemic waves is vital to epidemic and pandemic preparedness and response. For example, in the midst of an outbreak, it may allow public health officials to target interventions toward subpopulations still at risk for significant transmission. Identifying informative indicators of urban community structure, including the insularity estimates proposed in this study, will allow us to construct temporal epidemiological risk maps and target interventions (eg, vaccination) toward highly connected individuals reaching outside their own communities that can potentially serve as bridge superspreaders. In the Montreal network, communities had distinct geographic signatures, suggesting that investigations of community structure may also facilitate the effective placement of surveillance or intervention sites such as vaccination clinics.

Acknowledgments

This work was funded by NIH/NIGMS MIDAS grant U01 GM087719. AGH is supported by NIH/NLM career development award K01 LM011985. JSB is supported by NIH/NLM grant R01LM010812. The authors are grateful to Genevieve Bastien for assistance with the ÎSF database; to David L Buckeridge, Aman Verma, Ting Hu, and Christian Darabos for helpful discussions; and to 2 anonymous reviewers for their helpful comments. The authors also acknowledge the Texas Advanced Computing Center at The University of Texas at Austin for providing high performance computing resources that have contributed to the research results reported within this paper.

Conflicts of Interest

ML is a cofounder of the not-for-profit Île Sans Fil, a community wireless network in Montreal. The other authors have no conflicts to disclose.

Multimedia Appendix 1

Supplementary methods.

[PDF File (Adobe PDF File), 631KB - jmir_v17i7e169_app1.pdf]

Multimedia Appendix 2

Supplementary figures.

[PDF File (Adobe PDF File), 6MB - jmir_v17i7e169_app2.pdf]

XSL•FO
RenderX

References

1. Miller MA, Viboud C, Balinska M, Simonsen L. The signature features of influenza pandemics--implications for policy. N Engl J Med 2009 Jun 18;360(25):2595-2598. [doi: 10.1056/NEJMp0903906] [Medline: 19423872]
2. Hsieh YH, Chen CWS. Turning points, reproduction number, and impact of climatological events for multi-wave dengue outbreaks. Trop Med Int Health 2009 Jun;14(6):628-638 [FREE Full text] [doi: 10.1111/j.1365-3156.2009.02277.x] [Medline: 19392743]
3. Lipsitch M, Cohen T, Cooper B, Robins JM, Ma S, James L, et al. Transmission dynamics and control of severe acute respiratory syndrome. Science 2003 Jun 20;300(5627):1966-1970 [FREE Full text] [doi: 10.1126/science.1086616] [Medline: 12766207]
4. Simonsen L, Olson D, Viboud C, Heiman E. Pandemic influenza and mortality: past evidence and projections for the future. In: Knobler SL, Mack A, Mahmoud LS, editors. The Threat of Pandemic Influenza: Are We Ready?. Washington, DC: National Academies Press; 2005:89-114.
5. Johnson NP, Mueller J. Updating the accounts: global mortality of the 1918-1920 "Spanish" influenza pandemic. Bull Hist Med 2002;76(1):105-115. [Medline: 11875246]
6. Frost WH. Statistics of influenza morbidity: with special reference to certain factors in case incidence and case fatality. Public Health Rep 1920;35(11):584. [doi: 10.2307/4575511]
7. Ammon CE. Spanish flu epidemic in 1918 in Geneva, Switzerland. Euro Surveill 2002 Dec;7(12):190-192. [Medline: 12631986]
8. Andreasen V, Viboud C, Simonsen L. Epidemiologic characterization of the 1918 influenza pandemic summer wave in Copenhagen: implications for pandemic control strategies. J Infect Dis 2008 Jan 15;197(2):270-278 [FREE Full text] [doi: 10.1086/524065] [Medline: 18194088]
9. Rios-Doria D, Chowell G. Qualitative analysis of the level of cross-protection between epidemic waves of the 1918-1919 influenza pandemic. J Theor Biol 2009 Dec 21;261(4):584-592. [doi: 10.1016/j.jtbi.2009.08.020] [Medline: 19703472]
10. Barry JM, Viboud C, Simonsen L. Cross-protection between successive waves of the 1918-1919 influenza pandemic: epidemiological evidence from US Army camps and from Britain. J Infect Dis 2008 Nov 15;198(10):1427-1434 [FREE Full text] [doi: 10.1086/592454] [Medline: 18808337]
11. Shanks GD, Waller M, Mackenzie A, Brundage JF. Determinants of mortality in naval units during the 1918-19 influenza pandemic. Lancet Infect Dis 2011 Oct;11(10):793-799. [doi: 10.1016/S1473-3099(11)70151-7] [Medline: 21958582]
12. Dorigatti I, Cauchemez S, Ferguson NM. Increased transmissibility explains the third wave of infection by the 2009 H1N1 pandemic virus in England. Proc Natl Acad Sci U S A 2013 Aug 13;110(33):13422-13427 [FREE Full text] [doi: 10.1073/pnas.1303117110] [Medline: 23882078]
13. Bolton K, McCaw J, McVernon J, Mathews J. Exploring alternate immune hypotheses in dynamical models of the 1918-1919 influenza pandemic. Influenza Other Respir Viruses 2011;5(Suppl. 1):208-211.
14. Shaman J, Goldstein E, Lipsitch M. Absolute humidity and pandemic versus epidemic influenza. Am J Epidemiol 2011 Jan 15;173(2):127-135 [FREE Full text] [doi: 10.1093/aje/kwq347] [Medline: 21081646]
15. He D, Dushoff J, Day T, Ma J, Earn DJD. Inferring the causes of the three waves of the 1918 influenza pandemic in England and Wales. Proc Biol Sci 2013 Sep 7;280(1766):20131345 [FREE Full text] [doi: 10.1098/rspb.2013.1345] [Medline: 23843396]
16. Caley P, Philp DJ, McCracken K. Quantifying social distancing arising from pandemic influenza. J R Soc Interface 2008 Jun 6;5(23):631-639 [FREE Full text] [doi: 10.1098/rsif.2007.1197] [Medline: 17916550]
17. Bootsma MCJ, Ferguson NM. The effect of public health measures on the 1918 influenza pandemic in US cities. Proc Natl Acad Sci U S A 2007 May 1;104(18):7588-7593 [FREE Full text] [doi: 10.1073/pnas.0611071104] [Medline: 17416677]
18. Girvan M, Newman MEJ. Community structure in social and biological networks. Proc Natl Acad Sci U S A 2002 Jun 11;99(12):7821-7826 [FREE Full text] [doi: 10.1073/pnas.122653799] [Medline: 12060727]
19. Watts DJ, Muhamad R, Medina DC, Dodds PS. Multiscale, resurgent epidemics in a hierarchical metapopulation model. Proc Natl Acad Sci U S A 2005 Aug 9;102(32):11157-11162 [FREE Full text] [doi: 10.1073/pnas.0501226102] [Medline: 16055564]
20. Christakis NA, Fowler JH. Social network sensors for early detection of contagious outbreaks. PLoS One 2010;5(9):e12948 [FREE Full text] [doi: 10.1371/journal.pone.0012948] [Medline: 20856792]
21. Salathé M, Kazandjieva M, Lee JW, Levis P, Feldman MW, Jones JH. A high-resolution human contact network for infectious disease transmission. Proc Natl Acad Sci U S A 2010 Dec 21;107(51):22020-22025 [FREE Full text] [doi: 10.1073/pnas.1009094108] [Medline: 21149721]
22. Eagle N, Pentland AS, Lazer D. Inferring friendship network structure by using mobile phone data. Proc Natl Acad Sci U S A 2009 Sep 8;106(36):15274-15278 [FREE Full text] [doi: 10.1073/pnas.0900282106] [Medline: 19706491]
23. Sun L, Axhausen KW, Lee DH, Huang X. Understanding metropolitan patterns of daily encounters. Proc Natl Acad Sci U S A 2013 Aug 20;110(34):13774-13779 [FREE Full text] [doi: 10.1073/pnas.1306440110] [Medline: 23918373]
24. Salathé M, Jones JH. Dynamics and control of diseases in networks with community structure. PLoS Comput Biol 2010 Apr;6(4):e1000736 [FREE Full text] [doi: 10.1371/journal.pcbi.1000736] [Medline: 20386735]

XSL•FO
RenderX

25. Guimerà R, Danon L, Díaz-Guilera A, Giralt F, Arenas A. Self-similar community structure in a network of human interactions. Phys Rev E Stat Nonlin Soft Matter Phys 2003 Dec;68(6 Pt 2):065103. [Medline: 14754250]
26. Hébert-Dufresne L, Allard A, Young JG, Dubé LJ. Global efficiency of local immunization on complex networks. Sci Rep 2013;3:2171 [FREE Full text] [doi: 10.1038/srep02171] [Medline: 23842121]
27. CRAWDAD. URL: http://crawdad.cs.dartmouth.edu/ [accessed 2014-07-22] [WebCite Cache ID 6RGGsjddg]
28. Clauset A, Newman M, Moore C. Finding community structure in very large networks. Phys Rev E 2004 Dec;70(6):1-6. [doi: 10.1103/PhysRevE.70.066111]
29. Hladish T, Melamud E, Barrera LA, Galvani A, Meyers LA. EpiFire: An open source C++ library and application for contact network epidemiology. BMC Bioinformatics 2012;13:76 [FREE Full text] [doi: 10.1186/1471-2105-13-76] [Medline: 22559915]
30. Abbey H. An examination of the Reed-Frost theory of epidemics. Hum Biol 1952 Sep;24(3):201-233. [Medline: 12990130]
31. Ferguson NM, Cummings DAT, Fraser C, Cajka JC, Cooley PC, Burke DS. Strategies for mitigating an influenza pandemic. Nature 2006 Jul 27;442(7101):448-452. [doi: 10.1038/nature04795] [Medline: 16642006]
32. Fraser C, Donnelly CA, Cauchemez S, Hanage WP, Van Kerkhove MD, Hollingsworth TD, WHO Rapid Pandemic Assessment Collaboration. Pandemic potential of a strain of influenza A (H1N1): early findings. Science 2009 Jun 19;324(5934):1557-1561 [FREE Full text] [doi: 10.1126/science.1176062] [Medline: 19433588]
33. Cruz-Pacheco G, Duran L, Esteva L, Minzoni AA, Lopez-Cervantes M, Panayotaros P, et al. Modelling of the influenza A(H1N1)v outbreak in Mexico City, April-May 2009, with control sanitary measures. Euro Surveill 2009 Jul 2;14(26):344-358 [FREE Full text] [Medline: 19573510]
34. Mills CE, Robins JM, Lipsitch M. Transmissibility of 1918 pandemic influenza. Nature 2004 Dec 16;432(7019):904-906. [doi: 10.1038/nature03063] [Medline: 15602562]
35. Chowell G, Nishiura H, Bettencourt LMA. Comparative estimation of the reproduction number for pandemic influenza from daily case notification data. J R Soc Interface 2007 Feb 22;4(12):155-166 [FREE Full text] [doi: 10.1098/rsif.2006.0161] [Medline: 17254982]
36. Newman M. Networks: An Introduction. Oxford: Oxford University Press; 2010.
37. Anderson R, May R. Infectious Diseases of Humans: Dynamics and Control. Oxford: Oxford University Press; 1991.
38. Meyers LA, Pourbohloul B, Newman MEJ, Skowronski DM, Brunham RC. Network theory and SARS: predicting outbreak diversity. J Theor Biol 2005 Jan 7;232(1):71-81. [doi: 10.1016/j.jtbi.2004.07.026] [Medline: 15498594]
39. Newman MEJ. Spread of epidemic disease on networks. Phys Rev E Stat Nonlin Soft Matter Phys 2002 Jul;66(1 Pt 2):016128. [Medline: 12241447]
40. Ball F, Neal P. Network epidemic models with two levels of mixing. Math Biosci 2008 Mar;212(1):69-87. [doi: 10.1016/j.mbs.2008.01.001] [Medline: 18280521]
41. McDonald JC. Influenza in Canada. Can Med Assoc J 1967 Sep 2;97(10):522-527 [FREE Full text] [Medline: 6031153]
42. Launay E, Ovetchkine P, Saint-Jean M, Coïc L, Ducruet T, Charest H, et al. Novel influenza A (H1N1): clinical features of pediatric hospitalizations in two successive waves. Int J Infect Dis 2011 Feb;15(2):e122-e130 [FREE Full text] [doi: 10.1016/j.ijid.2010.08.006] [Medline: 21144785]
43. He D, Dushoff J, Day T, Ma J, Earn DJD. Mechanistic modelling of the three waves of the 1918 influenza pandemic. Theor Ecol 2011;4(2):283-288.
44. Cauchemez S, Valleron AJ, Boëlle PY, Flahault A, Ferguson NM. Estimating the impact of school closure on influenza transmission from Sentinel data. Nature 2008 Apr 10;452(7188):750-754. [doi: 10.1038/nature06732] [Medline: 18401408]

Abbreviations

CRAWDAD: Community Resource for Archiving Wireless Data at Dartmouth
ÎSF: Île Sans Fil
R_0: basic reproductive number

Edited by G Eysenbach; submitted 25.07.14; peer-reviewed by E Lau, A Hyder; comments to author 25.11.14; revised version received 23.01.15; accepted 23.03.15; published 08.07.15.

Please cite as:
Hoen AG, Hladish TJ, Eggo RM, Lenczner M, Brownstein JS, Meyers LA
Epidemic Wave Dynamics Attributable to Urban Community Structure: A Theoretical Characterization of Disease Transmission in a Large Network
J Med Internet Res 2015;17(7):e169
URL: http://www.jmir.org/2015/7/e169/
doi:10.2196/jmir.3720
PMID:26156032

XSL•FO
RenderX

©Anne G Hoen, Thomas J Hladish, Rosalind M Eggo, Michael Lenczner, John S Brownstein, Lauren Ancel Meyers. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 08.07.2015. 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

Characterizing the Processes for Navigating Internet Health Information Using Real-Time Observations: A Mixed-Methods Approach

Susan L Perez[1,2], MPH, PhD; Debora A Paterniti[1,3,4,5], PhD; Machelle Wilson[5], PhD; Robert A Bell[6,7], PhD; Man Shan Chan[1], BS; Chloe C Villareal[4], BA; Hien Huy Nguyen[3,5,8], MD; Richard L Kravitz[3], MD, MSPH

[1]Betty Irene Moore School of Nursing, University of California, Davis, Sacramento, CA, United States

[2]Department of Kinesiology and Health Science, California State University, Sacramento, Sacramento, CA, United States

[3]Department of Internal Medicine, University of California, Davis, Sacramento, CA, United States

[4]Department of Sociology, University of California, Davis, Davis, CA, United States

[5]Department of Public Health Sciences, Division of Biostatistics, University of California, Davis, Sacramento, CA, United States

[6]Department of Communication, University of California, Davis, Davis, CA, United States

[7]Department of Public Health Sciences, University of California, Davis, Davis, CA, United States

[8]Department of Infectious Disease, University of California, Davis, Sacramento, CA, United States

Corresponding Author:
Susan L Perez, MPH, PhD
Betty Irene Moore School of Nursing
University of California, Davis
2103 Stockton Blvd
Sacramento, CA, 95817
United States
Phone: 1 5308488011
Fax: 1 9167343257
Email: susan.perez@gmail.com

Abstract

Background: Little is known about the processes people use to find health-related information on the Internet or the individual characteristics that shape selection of information-seeking approaches.

Objective: Our aim was to describe the processes by which users navigate the Internet for information about a hypothetical acute illness and to identify individual characteristics predictive of their information-seeking strategies.

Methods: Study participants were recruited from public settings and agencies. Interested individuals were screened for eligibility using an online questionnaire. Participants listened to one of two clinical scenarios—consistent with influenza or bacterial meningitis—and then conducted an Internet search. Screen-capture video software captured Internet search mouse clicks and keystrokes. Each step of the search was coded as hypothesis testing (etiology), evidence gathering (symptoms), or action/treatment seeking (behavior). The coded steps were used to form a step-by-step pattern of each participant's information-seeking process. A total of 78 Internet health information seekers ranging from 21-35 years of age and who experienced barriers to accessing health care services participated.

Results: We identified 27 unique patterns of information seeking, which were grouped into four overarching classifications based on the number of steps taken during the search, whether a pattern consisted of developing a hypothesis and exploring symptoms before ending the search or searching an action/treatment, and whether a pattern ended with action/treatment seeking. Applying dual-processing theory, we categorized the four overarching pattern classifications as either System 1 (41%, 32/78), unconscious, rapid, automatic, and high capacity processing; or System 2 (59%, 46/78), conscious, slow, and deliberative processing. Using multivariate regression, we found that System 2 processing was associated with higher education and younger age.

Conclusions: We identified and classified two approaches to processing Internet health information. System 2 processing, a methodical approach, most resembles the strategies for information processing that have been found in other studies to be associated with higher-quality decisions. We conclude that the quality of Internet health-information seeking could be improved through

XSL•FO
RenderX

consumer education on methodical Internet navigation strategies and the incorporation of decision aids into health information websites.

(J Med Internet Res 2015;17(7):e173) doi:10.2196/jmir.3945

KEYWORDS

dual processing; information seeking; Internet search; health information

Introduction

The Internet has developed into a poorly organized information space of varying quality [1]. Its rapid growth has posed a serious problem due to people's limited cognitive abilities to process the masses of information encountered during a typical Internet search [2]. Attempts have been made to understand the ways in which health information seekers cope with the unruly structure of the Internet. For example, investigators have examined the general intuitive strategies information seekers use to process complex online information [3-5].

In addition, several studies have used observational and survey methods to better understand how people undertake Internet health information searches in response to specific health-related questions and situations [3,6,7]. Findings from these studies suggest that health information search processes vary depending on current health circumstances and previous health experiences. In addition to situational factors, such as topic familiarity and complexity, there is evidence of variation in search strategies based on individual characteristics, such as gender, insurance status, education, and age [8]. If search patterns vary systematically by demographic and personal characteristics, it may ultimately be feasible to create targeted content and delivery systems that match up with group-level needs and preferences [9].

Our study focuses on consumers' use of the Internet to interpret symptoms and reach a preliminary diagnosis. Such research is warranted by the fact that 35% of US adults use the Internet for self-diagnosis [7]. Specifically, we investigate the Internet health information search processes used to make health-related decisions amid the challenges that come with Internet navigation and the literacy levels required to decipher medical information [10].

Three main paradigms in psychology of judgment and decision making may inform how people seek information in response to acute, troubling symptoms [11]: (1) heuristics and biases research (also known as dual-processing theory) that focuses on an individual's judgment of probability [12], (2) the study of decision making under risk [13], and (3) social judgment theory (as applied in the lens model) [14]. This investigation is grounded in dual-processing theory [15,16] because this theory emphasizes judgment under uncertainty [12] such as what one may encounter when seeking information to inform a response to troubling medical symptoms.

Dual-processing theory posits that two distinct cognitive systems (System 1 and System 2) are invoked during human decision making [12,17]. System 1 processing triggers the use of biases and heuristics, while System 2 processing is a methodical evaluation of the information presented. The use of System 2 for information processing may reduce the impact of the intuitive biases in the automated processes, which makes for sounder decision making [18]. Laymen and experts alike are prone to invoke heuristic biases characteristic of System 1 processing when pressed to think intuitively [19]. In order to reduce the biases imposed by System 1 thinking in order to arrive at a high-quality health decision, it is important to find ways to enforce System 2 processing. Understanding how Internet searchers enact strategies calling upon System 1 and System 2 thinking could have important implications for the way in which information on health websites is organized and presented.

An understanding of the processes by which Internet users seek, attend to, and assimilate health information can help webpage developers anticipate user needs and guide their attention to materials that support higher quality decisions [1,4]—decisions that are consistent with evidence-based practices and the patient's own values. It may be particularly useful to understand when and how Internet health information seekers adopt System 1 versus System 2 processing. For all but the simplest decisions, the use of System 2 decreases bias and is associated with better decisions [18].

Additionally, an understanding of whether, how, and to what extent Internet users invoke information processing strategies during actual Internet health information searches could inform the design of Internet health search engines and websites [20]. Consider Internet users who invoke System 1 thinking when researching topics that are not familiar to them or topics that are complex. The introduction of features designed to slow down information-seeking processes, such as graphics, animations, sidebars, or quizzes, could prompt a shift towards System 2 thinking, leading to more deliberate, high-quality processing of information.

Previous studies have yet to characterize the processes by which individuals navigate Internet health information when making a health decision. We thus designed an observational study to explore strategies people use in seeking health-related information on the Internet and to understand the factors that predict their approach to finding and processing information. To standardize the stimuli provided to study participants, we used two clinical vignettes representing acute illnesses of contrasting clinical severity. There were no other constraints placed upon the participants as they investigated these two illnesses. In so doing, we aimed to describe Internet search processes and to identify demographic and personal characteristics associated with use of System 1 and System 2 cognitive processing.

XSL•FO
RenderX

Methods

Ethical Approval

The Institutional Review Board at the University of California, Davis, approved all study procedures.

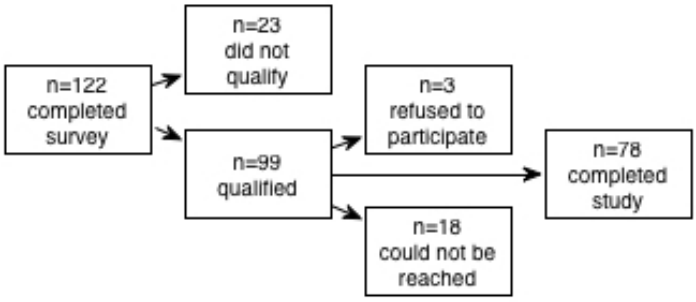
Recruitment of Participants

This investigation was designed as an exploratory, mixed-methods study focused on the Internet search patterns of young adults with limited access to a regular health care provider and services. Participant recruitment and data collection methods were adopted from previous studies [3,6]. We studied this age group because this demographic is more likely to experience barriers to health care continuity due to transitions in life and career [21], which may persist even with the implementation of improved coverage provisions such as the Affordable Care Act [22] and use of the Internet for health information [7,23]. The goal was to recruit at least 75 participants over a 6-month period. Participants were eligible if they were 21-35 years of age, had searched the Internet for health information within the past 12 months, and reported at least one of four barriers to accessing health care services (diminished ability to pay for services, no established relationship with a trusted primary care provider, an inability to get an appointment in a timely manner, or limited transportation options) [24].

Participants were recruited and data collected between March and August 2013. Potential participants were identified through door-to-door canvassing in a low-income housing community; canvassing at community fairs, community colleges, and local government offices that offer public services; sending emails through a University listserv targeting minority students; and posting flyers within the geographical region of Yolo County, California, at local coffee shops, public libraries, student family housing complexes, and community colleges. Individuals who expressed interest in participating were screened for eligibility using an online questionnaire (see Multimedia Appendix 1).

In total, 78 of the 122 people screened were eligible (see Figure 1). Those who qualified and agreed to participate were contacted to schedule an interview at one of four public libraries.

Figure 1. A breakdown of the number of individuals who completed the screening survey, qualified as eligible, agreed to participate, and completed the study.



Study Procedures

Data were collected individually from each participant. Upon arrival at the study site, the participant was accompanied by the lead author to a quiet room or cubicle where they were provided written informed consent, participated in a brief demographic questionnaire, assessed for health status using the Short-Form (SF)-36 Health Survey [25], and instructed on study procedures and oriented to a laptop computer. For practice, participants were first asked to participate in a "mock search" focused on purchasing a box of chocolates. While conducting the mock search, subjects were asked to describe the actions they were taking, the content they were reading, and the qualities of the webpages that drew their attention. This instruction also oriented the participant to the operation of the laptop computer used for this study. Following this training exercise, the participant was randomly assigned to one of two clinical symptom scenarios of varying severity involving (1) fever, mild headache, dry cough, and myalgia, suggestive of influenza, and (2) fever, severe headache, and stiff neck, suggestive of meningitis. As a prompt to the clinical symptoms scenario, each participant was instructed to "Imagine you are experiencing this situation or think about a time when you had experienced this situation". Unless the participant inquired, participants were not informed that the symptoms were suggestive of influenza or meningitis.

See Multimedia Appendix 2 for an example of an Internet health information search.

The symptom scenarios were developed based on Centers for Disease Control and Prevention guidelines with input from clinical co-authors (RLK, HN). Both symptom scenarios were pilot tested for comprehensibility in a small sample of young adults (n=8). A total of 42 of the 78 participants were randomly assigned to the influenza scenario and 36 to the meningitis scenario. The participant was then instructed to "search the Internet, as though you were experiencing this situation" and trained to "think out loud" while doing so. The participant had a choice of Web browser that included Firefox, Internet Explorer, or Google Chrome. All Web browsers opened to a blank page. Participants' Internet searches and think-out-loud vocalizations were digitally recorded using screen capture video-recording software [26]. Browser search history and cookies were deleted after each data collection session. Upon completion, the participant received a payment of US $20.

Data Preparation and Coding

The digital video recordings were transferred from the laptop computer used for data collection to a computer used for analysis and saved to electronic files. While reviewing the video recordings of the Internet search, team members developed a

XSL•FO
RenderX

list of interactions. "Interaction" was defined as the entering of a search term, selection of a website, and selection of a link to a website. We used these data to create a chronological workflow of each participant's Internet search interactions. Members of the research team (SLP, CCV, and MSC) developed specific codes to be applied to the interactions through a process of analytical induction. Team members met regularly to compare their coding of participant's interactions. Disagreements were resolved through discussion until members of the team reached consensus.

Coding of Internet Search Behaviors

Each interaction was classified as one of three search units (SU): (1) hypothesis testing, (2) evidence gathering, or (3) treatment/action seeking [27]. Hypothesis testing describes interactions relevant to testing a diagnostic hypothesis (ie, entering the search term "meningitis" or clicking on a hyperlink titled "Flu"). Evidence gathering describes interactions involving symptoms (eg, "achy, high temperature, sore muscles" or clicking on the link "cough, muscle pain / symptom search"). Treatment/action seeking describes interactions that address remedies, recommended actions, or alerts such as recommendations for seeking immediate care from a health care provider, looking for a cure, or searching for health care services (ie, entering the search term "flu remedies" or selecting the link "When to Seek Medical Care").

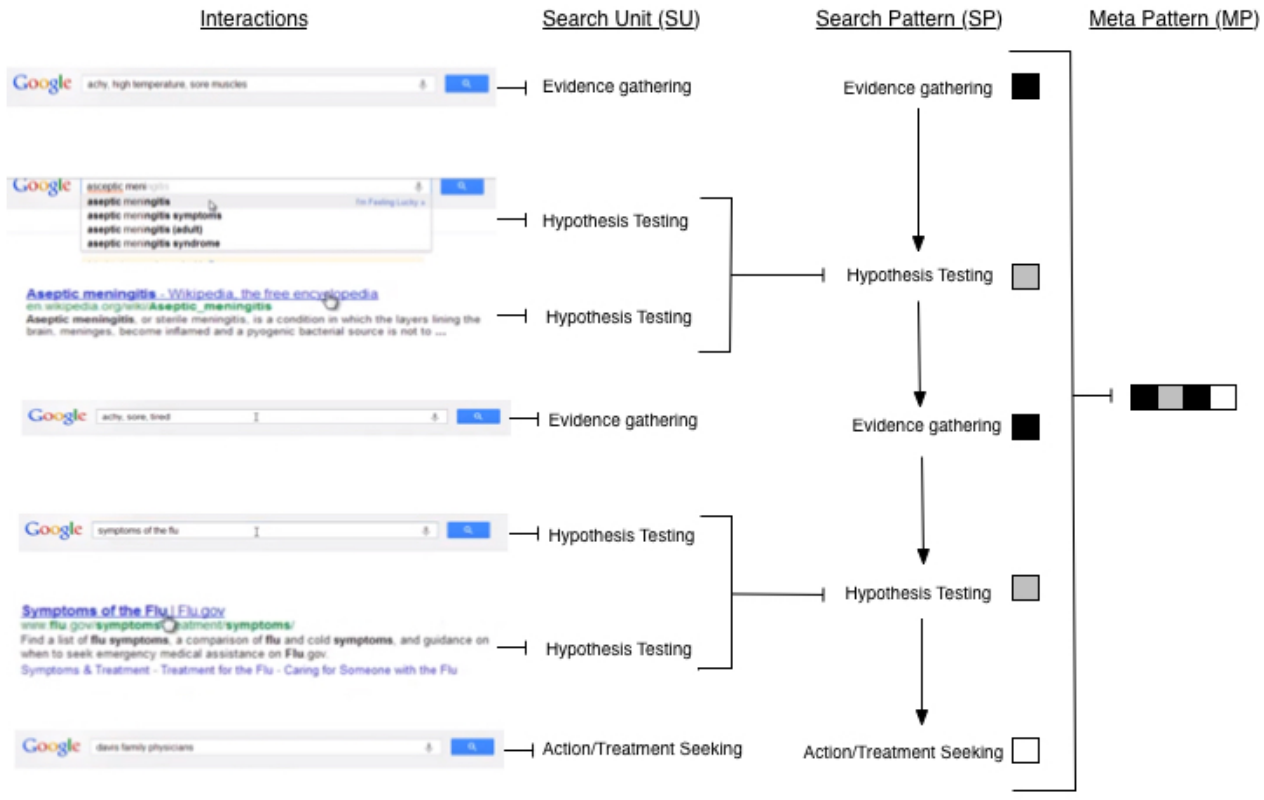
Next, unbroken sequences of one, two, or more identical SUs were deemed search patterns (SPs). SPs are higher-order categories consisting of one or more SUs. For example, if a participant entered a query for "asceptic meni", selected a link titled "aseptic meningitis", and selected a link titled "Aseptic meningitis – Wikipedia the free encyclopedia", these three consecutive hypothesis testing SUs would be merged to form one hypothesis testing SP as shown in the middle panel of Figure 2.

As the last step in the Internet search coding process, SPs for a given individual were ordered chronologically. The resulting sequences were displayed graphically and then categorized into a small set of overarching patterns called meta-patterns (MPs). Figure 2 provides an example of the procedure for converting Internet interactions into a MP. This example forms the following meta-pattern:

evidence gathering → hypothesis testing → evidence gathering → hypothesis testing → action/treatment seeking

The MPs were organized into four overarching pattern classifications using a hierarchical system for classification based on number of SPs (≥1 and <1), inclusion of a hypothesis testing and/or evidence gathering SP pattern combination (in no specific order), and inclusion of action/treatment seeking before termination of the search. Finally, the four overarching pattern classifications were examined in light of the dual-processing theory framework, allowing each Internet search to be categorized as predominantly System 1 or System 2. The patterns classified as including a hypothesis testing and evidence gathering SP combination (in no specific order) were categorized as System 2. All other pattern classifications were categorized as System 1.

Figure 2. Example of coding process from interaction to meta-pattern.



XSL•FO
RenderX

Statistical Analysis

We identified demographic and personal characteristic differences between respondents using System 1 processing and those using System 2. A *t* test was used to make this two-group comparison for age in years, as well as SF-36 scales assessing physical functioning, role-physical, bodily pain, general health, vitality, social functioning, role-emotional, and mental health [25]. The chi-square test was used for the categorical variables of race, gender, treatment, recruitment site, and education level. Respondent characteristics with a *P* value ≤.1 or less were included in a multivariate logistic regression model. In this analysis, dominant search strategy served as the dependent variable, with System 1 serving as the reference group. We fit a logistic regression model using a backward selection procedure to test for effects of the SF-36 scale physical functioning, site, gender, race, and education. A *P* value of .05 was considered significant. All statistical analyses were performed using SAS(r) software version 9.3.

Results

Sample Characteristics

A summary of participant demographic characteristics is presented in Table 1. The sample was young, with a mean age of 25 (SD 4.38), predominantly female, nonwhite with a strong Hispanic representation, not college educated, and health-insured. When comparing the average values of study participants' demographic characteristics with residents of Yolo County [28,29] as a whole, all the variables, except education, were significantly different (*P*>.05). In order to conserve power, the racial groups of Asian (4/78, 5%), American Indian or Alaska Native (2/78, 3%), Black/African American (3/78, 4%), mixed race (4/78, 5%), other (29/78, 37%), and those who declined to state (13/78, 17%) were categorized as "Other".

Table 1. Summary of participant demographic characteristics.

Category	Study participants (N=78) n (%)	All Yolo County residents[a] (N=200,849) %
Age in years, mean	25	30[c]
Gender		
Male	23 (29)	49
Female	55 (71)	51
Race		
White	23 (29)	76
Other	55 (71)	24
Education		
No bachelor's degree	52 (66)	62
Bachelor's degree or higher	26 (35)	38
Insurance status		
Uninsured	18 (23)	20[b]
Public insurance	11 (14)	19[b]
Other insurance	49 (63)	68[b]
Total		

[a]2010 US Census [28].

[b]Among individuals under 65 years old; 2005 California Health Interview Survey [29].

[c]City-data.com.

Internet Search Patterns

The duration of Internet searching ranged from 0.92 minutes to 14.27 minutes with an average of 5.13 minutes. There was a moderate positive correlation between the number of interactions (eg, mouse clicks and entering search terms) and duration of Internet searching: r=.38, P<.001. There was wide variation in the number and ordering of search patterns (SPs). We identified 27 unique pattern variations of MPs as shown in Figure 3.

Working by consensus of the co-authors, we identified four overarching pattern classifications as depicted in Figure 4. These four overarching pattern classifications were created by first grouping search patterns based on number of steps taken during the search. A *simple* search was one involving a single step (21/78, 27%). A compound search was any search involving two or more steps (57/78, 73%).

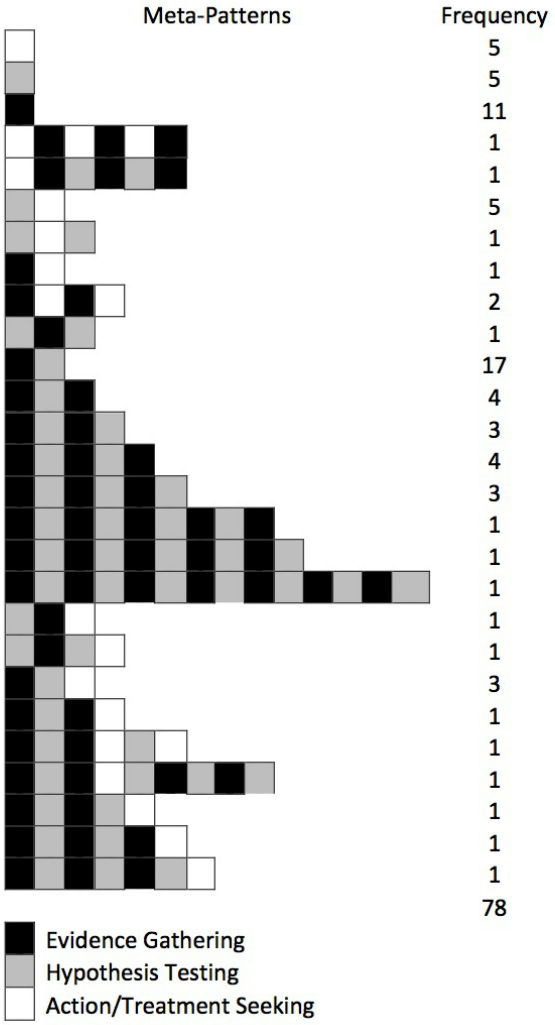
Thereafter, each compound pattern was placed into one of two subgroups, labeled *intuitive* and *analytical*. An intuitive search was any search that involved action/treatment seeking before hypothesis testing and evidence gathering were carried out. An analytical search was one that began with hypothesis testing and evidence gathering. An analytical search that did not lead

XSL•FO
RenderX

to action/treatment seeking was further classified as *analytical-recursive*. An analytical search that lead to action/treatment seeking was classified as *analytical-methodical*. Of the 57 compound searches, 11 (19%) were intuitive, 31 (54%) were analytical-recursive, and 15 (26%) were analytical-methodical.

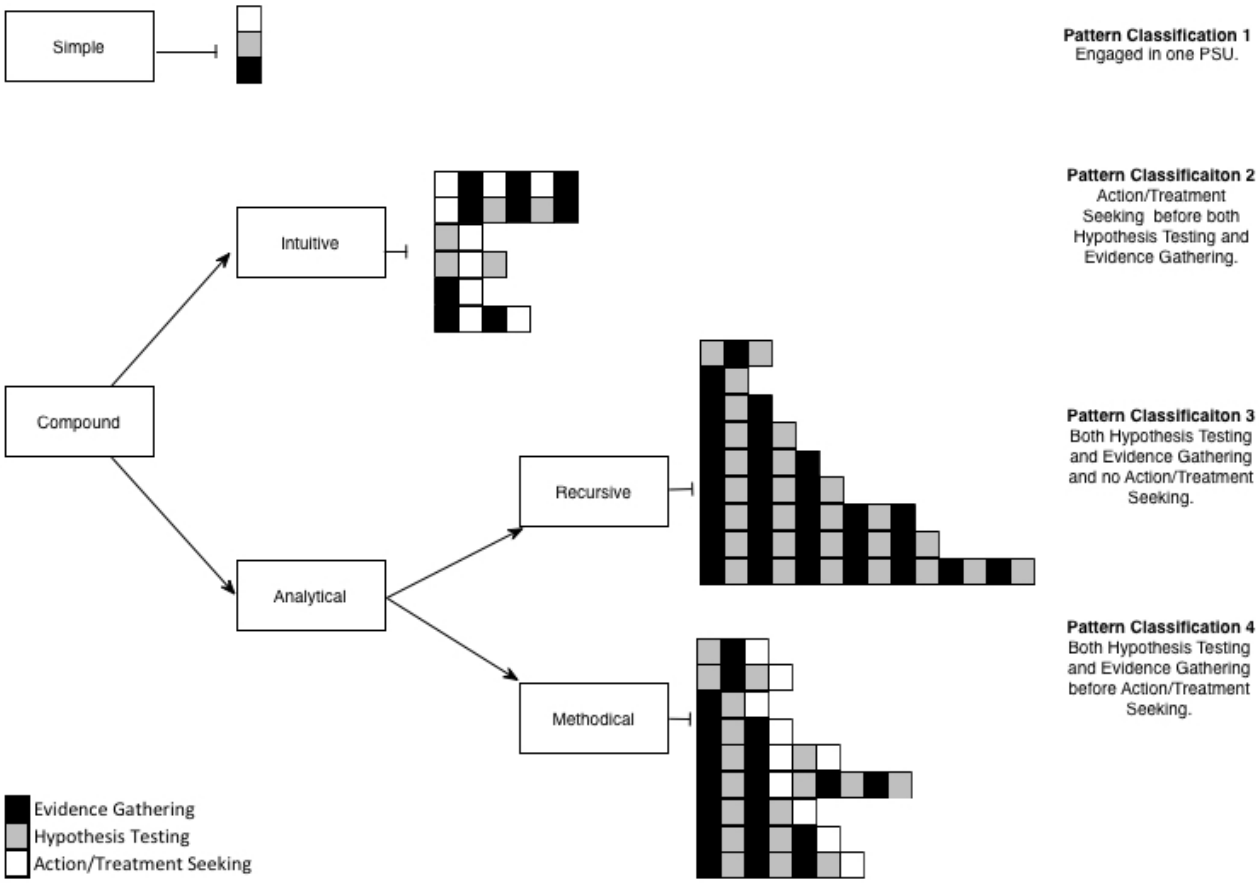
The four overarching pattern classifications we derived inductively fit neatly within the dual-processing theory for information processing. Pattern Classifications 1 and 2 were characterized as *System 1 processing* because these search patterns involved rapid progression to action/treatment or terminated after limited Internet searching, implying reliance on heuristic cues or ready satisfaction with Internet health information. Pattern Classifications 3 and 4 were characterized as *System 2 processing* (systematic) because these patterns involved a systematic approach to searching for information about specific diagnoses and symptoms prior to a search for actions/treatments or search termination. Of the 78 respondents, 32 (41%) engaged in System 1 processing and 46 (59%) relied on System 2 processing.

Figure 3. Frequency of each meta-pattern observed.



XSL•FO
RenderX

Figure 4. Classification of search strategies.



Predictors of Systematic Searching

A backward stepwise binary logistic regression model was constructed to predict a systematic approach to searching Internet health information (compared with System 1 processing) using age, clinical symptom scenario (influenza vs meningitis), male sex, white race, college education, and recruitment site not located within a university town as predictors (independent variables). The resulting model (Table 2) revealed a strong association between choice of systematic processing with education and age. Systematic processing was not significantly associated with symptom scenario, gender, race, or insurance status. For every 1-year increase in age, the odds of systematic processing decreased by 13.3% (OR 0.87, 95% CI 0.77-0.98, *P*=.02). There was also a significant negative effect of education on central route processing (OR 0.30, 95% CI 0.09-0.94, *P*=.04). Less educated participants (those without a bachelor's degree) were less likely to use systematic processing. There also were significant effects of recruitment site.

Table 2. Binary logistic regression analysis of the relationship of demographic characteristics with reliance on systematic search strategies (N=78).

Variable	OR	95% CI	*P*
Age	0.87	0.77-0.98	.02
Education (no bachelor's degree)	0.30	0.09-0.94	.04
Recruitment site	3.75	1.29-10.92	.02

Discussion

Principal Findings

In this study, we directly observed young adults as they searched for information on two hypothetical clinical scenarios varying in severity, influenza, and meningitis. The four overarching pattern classifications were categorized into two information-processing strategies as postulated by dual-process theory. While the results demonstrate a modest preference for behaviors associated with System 2 thinking, a substantial proportion of study participants relied on simple or "intuitive" approaches associated with System 1. In terms of predictors, we found that younger participants and those with more education were more likely embrace a System 2 approach.

Previous research on Internet health information seeking has focused on Internet information accuracy [10,30-36], completeness [30,32,36,37], and readability by the lay public [10,38]. Other investigators have focused on user information preferences and needs [3], demographics of individuals with specific information preferences [7], and user accounts of preferences for information format [3]. These studies, evaluating Internet health information and user characteristics, suggest that

XSL•FO
RenderX

Internet health information is often inaccurate, incomplete, and difficult to comprehend by lay audiences, yet individuals are still turning to the Internet for health decision making [7].

Little can be done to address the inconsistencies of the millions of Internet health information websites [39], but we can find ways to guide Internet health information seekers toward information processing strategies that may be more likely to lead to accurate decision making. Prior research concerning the design of information systems indicates that to support accurate decision making, systems should be process-oriented rather than information-oriented because information seeking and decision making involve a series of encounters over time rather than a single information encounter [40,41]. The distinction between automatic (System 1) and controlled (System 2) processing explained the roles of motivation and cognitive ability in the decision-making processes of the young adults studies [20].

The decision-making process most conducive to a high-quality decision involves systematically gathering all available information about a situation, weighing every feasible option, and integrating the available data to make the decision most likely to produce desired outcomes [42]. When there is no single "best" action, a high-quality decision balances the subjective values of the consumer's assessment of benefits versus harms [43]. Of the two observed approaches to Internet health information seeking, the behaviors associated with System 2 may be most conducive to a high-quality decision because System 2 processors methodically develop a hypothesis (eg, a provisional conjecture established from information gathered during the Internet search or previously held knowledge) and gather information (eg, gathering information to confirm or develop a hypothesis) before taking action/treating or terminating their search.

System 1 health information seekers are more likely to reach decisions based on simplifying heuristic rules and to terminate their search once they have found an acceptable solution, not necessarily the best. The System 1 approach is often effective because it economizes on time and usually leads to sensible decisions [42]. However, departures from the ideal strategy of information seeking may lead to mistakes, such as errors in reasoning that arise from misinformation, denial, overconfidence, distrust, or confusion [42].

Limitations

This study has limitations. First, generalizability is limited by a small sample size of young adults recruited by convenience sampling from a limited geographical region in central California. Second, the participant characteristics of gender, race, and insurance status did not fully reflect the demographics of the region from which participants were recruited. Third, participants may have had more or less familiarity with the symptom scenarios, which could bias their search process as a result of previous clinical experiences. Fourth, the unnatural and forced environment of our experiment may have influenced how subjects searched. Fifth, we likely omitted variables or factors that would potentially affect Internet search patterns. Finally, given their exploratory nature, these findings need to be validated with fresh samples.

Due to the unregulated nature of the Web itself, Internet health information seekers are susceptible to a wayward and distracted process of information gathering. Methodical Internet searching to guide health information seekers toward high-quality decisions should be approached in two ways: (1) through consumer education on methodical Internet navigation strategies and (2) through the incorporation of decision aids into health information websites.

Consumer education on a methodical approach to decision making when using the Internet would include guidance on the process of first defining the decision, then gathering information about potential decision outcomes, and ensuring that the final decision is consistent with the consumer's values [43]. The incorporation of decision aids into Internet health information websites should consider (1) defining the decision (health information resources should provide information about all options), (2) providing information about potential decision outcomes (health information resources should present probabilities, balance the presentation of options, and base information on up-to-date scientific evidence), and (3) supporting decision making that is consistent with consumers' values (health information resources should clarify and express values, use patient stories, and guide the deliberation process) [44].

Further research is needed to confirm the information-seeking processes most conducive to supporting high-quality decisions leading to the best possible outcomes. Professional health care providers can do little to control the type of health information encountered on the Internet, but they can help steer their patients towards Internet resources that encourage deliberative thinking and thus better decision making.

Acknowledgments

This project was supported in part by the National Center for Advancing Translational Sciences (NCATS), National Institutes of Health (NIH), through grant #UL1 TR000002.

Support for this publication was provided by grant number T32HS022236 from the Agency for Healthcare Research and Quality (AHRQ) through the Quality, Safety, and Comparative Effectiveness Research Training (QSCERT) Program.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1

Online questionnaire used to screen for eligibility to participate.

[PDF File (Adobe PDF File), 228KB - jmir_v17i7e173_app1.pdf]

Multimedia Appendix 2

Example of an Internet health information search, with annotations, in response to a vignette of symptoms mimicking meningitis.

[MP4 File (MP4 Video), 11MB - jmir_v17i7e173_app2.mp4]

References

1. Hölscher C, Strube G. Web search behavior of Internet experts and newbies. Computer Networks 2000 Jun;33(1-6):337-346. [doi: 10.1016/S1389-1286(00)00031-1]
2. Goldsmith J. How will the Internet change our health system? Health Affairs 2000 Jan 01;19(1):148-156. [doi: 10.1377/hlthaff.19.1.148]
3. 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]
4. Horvitz E, Kadie C, Paek T, Hovel D. Models of attention in computing and communication. Commun. ACM 2003 Mar 01;46(3):52. [doi: 10.1145/636772.636798]
5. Gil M, Serral E, Valderas P, Pelechano V. Designing for user attention: A method for supporting unobtrusive routine tasks. Science of Computer Programming 2013 Oct;78(10):1987-2008. [doi: 10.1016/j.scico.2013.03.002]
6. van Deursen AJAM, van Dijk JAGM. Internet skills performance tests: are people ready for eHealth? J Med Internet Res 2011;13(2):e35 [FREE Full text] [doi: 10.2196/jmir.1581] [Medline: 21531690]
7. Fox S, Duggan M. Pew Internet and American Life Project. Washington, DC: Pew Research Center's Internet and American Life Project; 2013. Health Online 2013 URL: http://www.pewinternet.org/files/old-media//Files/Reports/PIP_HealthOnline.pdf [accessed 2015-06-30] [WebCite Cache ID 6ZfNtj9v5]
8. Anker AE, Reinhart AM, Feeley TH. Health information seeking: a review of measures and methods. Patient Educ Couns 2011 Mar;82(3):346-354. [doi: 10.1016/j.pec.2010.12.008] [Medline: 21239134]
9. Suggs LS, McIntyre C. Are we there yet? An examination of online tailored health communication. Health Educ Behav 2009 Apr;36(2):278-288. [doi: 10.1177/1090198107303309] [Medline: 17620667]
10. Berland G, Elliott M, Morales L, Algazy J, Kravitz R, Broder M, et al. Health Information on the Internet. JAMA 2001 May 23;285(20):2612. [doi: 10.1001/jama.285.20.2612]
11. Evans JSBT. Dual-processing accounts of reasoning, judgment, and social cognition. Annu Rev Psychol 2008;59:255-278. [doi: 10.1146/annurev.psych.59.103006.093629] [Medline: 18154502]
12. Holyoak K, Morrison R, Frederick S. A Model of Heuristic Judgment. In: Holyoah K, Morrison R, editors. The Cambridge handbook of thinking and reasoning. New York: Cambridge University Press; 2005.
13. Harvey N, Koehler DJ, Zhang J, Gonzalez R. Decision under risk. In: Koehler D, editor. Blackwell handbook of judgment and decision making. Oxford, UK: Blackwell Pub; 2004.
14. Johnson W, Doherty M. Social judgment theory and academic advisement. Journal of Counseling Psychology 1983;30(2):271-274. [doi: 10.1037/0022-0167.30.2.271]
15. Chaiken S. Heuristic versus systematic information processing and the use of source versus message cues in persuasion. Journal of Personality and Social Psychology 1980;39(5):752-766. [doi: 10.1037/0022-3514.39.5.752]
16. Petty R, Cacioppo JT. Communication and Persuasion: Central and Peripheral Routes to Attitude Change. New York: Springer; 1986.
17. World Health Organization. Control of Health Care Costs in Social Security Systems. In: Report on a Workshop.: Regional Office for Europe, World Health Organization; 1982 Presented at: World Health Organization; May 25-28, 1981; Vienna.
18. Kahneman D, Lovallo D, Sibony O. Before you make that big decision. Harv Bus Rev 2011 Jun;89(6):50-60, 137. [Medline: 21714386]
19. Tversky A, Kahneman D. Judgment under Uncertainty: Heuristics and Biases. Science 1974 Sep 27;185(4157):1124-1131. [doi: 10.1126/science.185.4157.1124] [Medline: 17835457]
20. Metzger M. Making sense of credibility on the Web: Models for evaluating online information and recommendations for future research. J Am Soc Inf Sci 2007 Nov;58(13):2078-2091. [doi: 10.1002/asi.20672]
21. Schwartz K, Sommers BD. ASPE Research Brief. Washington, DC: Office of the Assistant Secretary for Planning and Evaluation, Department of Health and Human Services; 2012 Mar. Young Adults Are Particularly Likely to Gain Stable Health Insurance Coverage as a Result of the Affordable Care Act URL: http://aspe.hhs.gov/health/reports/2012/UninsuredYoungAdults/rb.shtml [accessed 2015-06-30] [WebCite Cache ID 6ZfOPljZ6]

XSL•FO
RenderX

22. Shartzer A, Long SK, Benatar S. Health Reform Monitoring Survey. Washington, DC: Urban Institute Health Policy Center; 2015 Jan 12. Health Care Costs Are a Barrier to Care for Many Women URL: http://hrms.urban.org/briefs/Health-Care-Costs-Are-a-Barrier-to-Care-for-Many-Women.html [accessed 2015-06-30] [WebCite Cache ID 6ZfOhP3vQ]
23. Fronstin P. California Healthcare Foundation. Sacramento, CA: California Healthcare Foundation; 2013. California's Uninsured: By the Numbers URL: http://www.chcf.org/publications/2013/12/californias-uninsured [accessed 2015-06-30] [WebCite Cache ID 6ZfOtHwFz]
24. Health People 2020. Washington, DC: Office of Disease Prevention and Health Promotion; 2015 Jun 29. Access to Health Services URL: http://www.healthypeople.gov/2020/topics-objectives/topic/Access-to-Health-Services [accessed 2015-06-30] [WebCite Cache ID 6ZfPChcgL]
25. Ware JE, Gandek B. Overview of the SF-36 Health Survey and the International Quality of Life Assessment (IQOLA) Project. J Clin Epidemiol 1998 Nov;51(11):903-912. [Medline: 9817107]
26. Camtasia Studio 8. Okemos, MI: TechSmith Corporation; 2015. URL: https://www.techsmith.com/camtasia.html [accessed 2015-06-30] [WebCite Cache ID 6ZfPl8lIy]
27. Kassirer J, Wong JK, Kopelman RE. Learning Clinical Reasoning. Baltimore, MD: Lippincott Williams & Wilkins; Sep 11, 2009.
28. State and County QuickFacts. 2010. United States Census Bureau URL: http://quickfacts.census.gov/qfd/states/06000.html [accessed 2015-06-30] [WebCite Cache ID 6ZfPWx2hm]
29. California Health Interview Survey. Los Angeles, CA: UCLA Center for Healthcare Policy and Research; 2005. CHIS Questionnaires URL: http://healthpolicy.ucla.edu/chis/design/Pages/questionnairesEnglish.aspx [accessed 2015-06-30] [WebCite Cache ID 6ZfPZ58nS]
30. Impicciatore P, Pandolfini C, Casella N, Bonati M. Reliability of health information for the public on the World Wide Web: systematic survey of advice on managing fever in children at home. BMJ 1997 Jun 28;314(7098):1875-1879 [FREE Full text] [Medline: 9224132]
31. Jejurikar SS, Rovak JM, Kuzon WM, Chung KC, Kotsis SV, Cederna PS. Evaluation of plastic surgery information on the Internet. Ann Plast Surg 2002 Nov;49(5):460-465. [doi: 10.1097/01.SAP.0000020053.67167.95] [Medline: 12439011]
32. Kulasegarah J. The quality of information on three common ENT procedures on the Internet. Ir J Med Sci 2007:221-224. [Medline: 22200968]
33. Kunst H, Groot D, Latthe PM, Latthe M, Khan KS. Accuracy of information on apparently credible websites: survey of five common health topics. BMJ 2002 Mar 9;324(7337):581-582 [FREE Full text] [Medline: 11884323]
34. Quinn E, Corrigan MA, McHugh SM, Murphy D, O'Mullane J, Hill ADK, et al. Breast cancer information on the internet: analysis of accessibility and accuracy. Breast 2012 Aug;21(4):514-517. [doi: 10.1016/j.breast.2012.01.020] [Medline: 22349349]
35. Scullard P, Peacock C, Davies P. Googling children's health: reliability of medical advice on the internet. Arch Dis Child 2010 Aug;95(8):580-582. [doi: 10.1136/adc.2009.168856] [Medline: 20371593]
36. Zun LS, Downey L, Brown S. Completeness and accuracy of emergency medical information on the web: update 2008. West J Emerg Med 2011 Nov;12(4):448-454 [FREE Full text] [doi: 10.5811/westjem.2010.10.1607] [Medline: 22224136]
37. Frické M, Fallis D, Jones M, Luszko GM. Consumer health information on the Internet about carpal tunnel syndrome: indicators of accuracy. Am J Med 2005 Feb;118(2):168-174. [doi: 10.1016/j.amjmed.2004.04.032] [Medline: 15694903]
38. Hendrick P, Ahmed OH, Bankier SS, Chan TJ, Crawford SA, Ryder CR, et al. Acute low back pain information online: an evaluation of quality, content accuracy and readability of related websites. Man Ther 2012 Aug;17(4):318-324. [doi: 10.1016/j.math.2012.02.019] [Medline: 22464886]
39. Kontos E, Bennett GG, Viswanath K. Barriers and facilitators to home computer and internet use among urban novice computer users of low socioeconomic position. J Med Internet Res 2007;9(4):e31 [FREE Full text] [doi: 10.2196/jmir.9.4.e31] [Medline: 17951215]
40. Soergel D. Organizing information: principles of data base and retrieval systems. Orlando, Fla: Academic Press; 1985.
41. Kuhlthau C. Inside the Search Process: Information Seeking from the User Perspective. Journal of the American Society for Information Science 1991;42(5):361-371. [doi: 10.1002/1097-4571]
42. Redelmeier DA, Rozin P, Kahneman D. Understanding patients' decisions. Cognitive and emotional perspectives. JAMA 1993 Jul 7;270(1):72-76. [Medline: 8510300]
43. Stacey D, Legare F, Nananda FC, Bennett CL. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2014;1(1):1-331. [doi: 10.1002/14651858.CD001431.pub4]
44. Elwyn G, O'Connor A, Stacey D, Volk R, Edwards A, Coulter A, International Patient Decision Aids Standards (IPDAS) Collaboration. Developing a quality criteria framework for patient decision aids: online international Delphi consensus process. BMJ 2006 Aug 26;333(7565):417 [FREE Full text] [doi: 10.1136/bmj.38926.629329.AE] [Medline: 16908462]

Abbreviations

MP: meta-patterns
SF-36: Short-Form-36 Health Survey

XSL•FO
RenderX

SP: search patterns
SU: search units

Edited by G Eysenbach; submitted 14.10.14; peer-reviewed by RS Sadasivam, P Ghezzi; comments to author 23.03.15; revised version received 01.05.15; accepted 24.05.15; published 20.07.15.

Please cite as:
Perez SL, Paterniti DA, Wilson M, Bell RA, Chan MS, Villareal CC, Nguyen HH, Kravitz RL
Characterizing the Processes for Navigating Internet Health Information Using Real-Time Observations: A Mixed-Methods Approach
J Med Internet Res 2015;17(7):e173
URL: http://www.jmir.org/2015/7/e173/
doi:10.2196/jmir.3945
PMID:26194787

©Susan L Perez, Debora A Paterniti, Machelle Wilson, Robert A Bell, Man Shan Chan, Chloe C Villareal, Hien Huy Nguyen, Richard L Kravitz. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 20.07.2015. 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

Secure Cloud-Based Solutions for Different eHealth Services in Spanish Rural Health Centers

Isabel de la Torre- Díez[1*], M Paed De, PhD (HK); Miguel Lopez-Coronado[1*], PhD (HK), M Paed De; Begonya Garcia-Zapirain Soto[2*], M Paed De, PhD (HK); Amaia Mendez-Zorrilla[2*], PhD (HK), M Paed De

[1]Grupo de Telemedicina y eSalud, Departamento de Teoría de la Señal y Comunicaciones e Ingeniería Telemática, University of Valladolid, Valladolid, Spain

[2]DeustoTech Institute of Technology, DeustoTech-Life Lab, University of Deusto, Bilbao, Spain

[*]all authors contributed equally

Corresponding Author:
Begonya Garcia-Zapirain Soto, M Paed De, PhD (HK)
DeustoTech Institute of Technology
DeustoTech-Life Lab
University of Deusto
Avda Universidades, 24
Bilbao, 48007
Spain
Phone: 34 944139000
Fax: 34 944139000
Email: mbgarciazapi@deusto.es

Abstract

Background: The combination of eHealth applications and/or services with cloud technology provides health care staff—with sufficient mobility and accessibility for them—to be able to transparently check any data they may need without having to worry about its physical location.

Objective: The main aim of this paper is to put forward secure cloud-based solutions for a range of eHealth services such as electronic health records (EHRs), telecardiology, teleconsultation, and telediagnosis.

Methods: The scenario chosen for introducing the services is a set of four rural health centers located within the same Spanish region. iCanCloud software was used to perform simulations in the proposed scenario. We chose online traffic and the cost per unit in terms of time as the parameters for choosing the secure solution on the most optimum cloud for each service.

Results: We suggest that load balancers always be fitted for all solutions in communication together with several Internet service providers and that smartcards be used to maintain identity to an appropriate extent. The solutions offered via private cloud for EHRs, teleconsultation, and telediagnosis services require a volume of online traffic calculated at being able to reach 2 Gbps per consultation. This may entail an average cost of €500/month.

Conclusions: The security solutions put forward for each eHealth service constitute an attempt to centralize all information on the cloud, thus offering greater accessibility to medical information in the case of EHRs alongside more reliable diagnoses and treatment for telecardiology, telediagnosis, and teleconsultation services. Therefore, better health care for the rural patient can be obtained at a reasonable cost.

(J Med Internet Res 2015;17(7):e157) doi:10.2196/jmir.4422

KEYWORDS

cloud; eHealth services; rural; security

Introduction

Background

Taking into account the economic, social, technological, and cultural transformations attached to the term “information society”, we as citizens, professionals, and lawmakers need to reconsider our standpoint regarding the population’s health requirements. Scientific advances, improvements in diagnosis, and therapeutic measures, in addition to a healthier lifestyle, have helped to add years to the life of the population. Within

XSL•FO
RenderX

this context, the emergence of telemedicine and telecare therefore becomes increasingly relevant. In modern day society, it is essential for doctors and other professionals to be able to provide their customers (patients) with the best possible information about the illness they have. Time, distance, and physical hindrances can no longer justify the barrier between the patient's illness and the best way of combating it, regardless of where one needs to go to obtain the solution. It is in such cases where telemedicine takes on more relevance by facilitating the use of expert advice.

Telemedicine can and must play a major role in situations where urgency, geography (rural areas), and other conditions (isolation, bad weather, catastrophes) require the use of this new health care model. This is similarly required in cases where towns, villages, or regions lack minimum services due to a critical lack of staff, such as in developing countries.

However, as telemedicine evolves, the need for certain basic principles that are accepted when put into practice becomes more apparent. We can view eHealth as a form of health practice backed up by electronic processes and information and communication technologies (ICT). Telecardiology, telediagnosis, teleconsultation, and electronic health records (EHRs) therefore constitute eHealth services devoted to improving the quality of a patient's treatment in terms of ensuring availability of specialists and reducing the need to travel, as well as being able to ensure a swift diagnosis and a second expert opinion.

The possibility of virtualizing resources on a cloud [1,2] will provide health care staff with sufficient mobility and accessibility to be able to transparently check any data they may need without having to worry about its physical location—they can still perform the tasks they deem appropriate using the information they require at any time [3-13]. Given that the data will not physically be in the same place as from where access is gained, special attention needs to be paid to how such services are requested while still ensuring optimum levels of security and privacy [14-17].

The combination of cloud computing with eHealth services can provide us with a virtual view of resources regardless of the geographic location or physical space they may occupy [6]. Fernández-Cardeñosa et al proposed two examples of cloud-based solutions for EHRs. One of them applied to a large hospital and the other one to primary care centers [18]. Although they conducted an economic analysis of the solutions, they failed to specify security aspects in the course of their work. Rodrigues et al later analyzed the security requirements of EHR solutions on the cloud [19]. They put forward secure and robust cloud-based solutions for equipping a set of rural health centers near Valladolid, Spain (where the hospital is located) with eHealth services such as EHRs, telecardiology, teleconsultation, and telediagnosis. These solutions may serve as a model when implementing eHealth services in other centers in other regions. In addition to providing models for introducing such secure solutions, we explain the importance of security within the framework of eHealth and any possible security frontiers.

In this paper, we discuss the security provided on the cloud within the framework of eHealth and its frontiers, the scenario chosen for the introduction of solutions, the authentication common to all solutions, and each eHealth solution (ie, EHRs on the cloud, telecardiology, teleconsultation, and telediagnosis).

Although the cloud offers obvious advantages, it can equally generate major concerns. When a network structure is run, it is exposed to denial-of-service attacks. If a user were able to take control of a service provider, that user would then be able to cease online services, with an ensuing cost of putting them in operation again. To put a stop to such attacks, the use of synchronized cookies as well as establishing a limit on the number of connected users help to neutralize Distributed Denial of Service (DDoS) attacks [20]. Another attack that the cloud may experience is in the case where, if the Secure Sockets Layer (SSL) is incorrectly configured, customer authentication will not be as expected, thus giving rise to a breach in security.

Therefore, the matter of security in cloud computing is a determining factor for this technology and one of the factors that can jeopardize its smooth running. How it inserts the user's details and operates with the user's programs—all within a physical place that does not belong to the user—may prove to be extremely discouraging for many. Well-known fraudulent methods used to obtain information such as phishing and botnets, etc, constitute serious threats to an organization's data and software [8]. One of the greatest concerns is then how to store data on the cloud—data should be transferred and stored in coded format by using proxies and agents to isolate the customer from direct access to shared storage on the cloud. Registers, audits, and compliance with regulations are features that require planning in cloud computing systems, and the concept of presence linked to identity must also be taken into account. The Internet is shown to be a flexible yet none-too-secure network—the greater the distributed system, the greater the possibility of attacks. Cloud computing is subject to the same vulnerabilities as Internet applications, plus others that may emerge from virtualizing and sharing resources and additional subcontracted services [21,22].

To assess the risks of implementation on the cloud, we need to conduct an analysis to determine which risks may be sensitive on our cloud and which mechanisms are to be used, etc. It is highly advisable for us to have an image of security that we are able to analyze in the search for vulnerabilities and dangers, so that we when we have doubts about the system's reliability, the only thing we need to do is return to that secure image.

Regarding security frontiers, the hardware/software bundle for the Cloud Security Alliance (CSA) model is taken as a reference for the cloud where the lowest level is Infrastructure as Software (IaaS), with Platform as Software (PaaS) and Service as Software (SaaS) above this. As levels increase in the bundle levels, each service model inherits the features of the one below—advantages and disadvantages as well as possible risks. IaaS offers the infrastructure, PaaS adds frameworks for developing applications, transactions and control structures, and SaaS is an operative environment with user applications, administration, and interface. From bottom to top, IaaS has the lowest functional and integrated security levels, and Saas the highest. Therefore, depending on which level we find ourselves, we will have certain security features or others, as well as

XSL•FO
RenderX

separating the responsibilities attached to the different levels of service model according to frontiers [22,23].

In the case of the SaaS model, the supplier offers security as part of the agreement regarding level of services, in which levels of compliance, Ugovernance, and responsibility for the whole bundle are stipulated. In the case of the PaaS mode, the security model can be defined in such a way that the supplier includes the software framework and a middleware layer—with this model, the customer would be responsible for the security of the application and the user interface on the upper part of the bundle. Last, the model with the lowest level of integrated security is IaaS, in which anything involving software of any type is the customer's problem.

Case Study Scenarios

The scenario proposed for offering secure eHealth solutions on the cloud is a set of rural health centers that provide different rural areas with basic care, with the most serious cases being those that require specialist care at the closest hospital—in this case, the one in Valladolid. The municipalities chosen are Peñafiel, Cuéllar, Tudela de Duero, and Portillo. The area has 36,000 inhabitants in total, each of which is covered by one doctor and one nurse who provide a physical presence. In addition, these health centers cover small neighboring villages, meaning that there is a variable flow of patients, and therefore numbers cannot be calculated exactly. These four health centers were chosen because of their proximity to the hospital in the regional capital and because they are the municipalities with the greatest number of patients referred to the cardiology department of the specific hospital (around 8% of hospital referrals). By way of an example of a rural health center, the town of Peñafiel in the province of Valladolid, Spain, was chosen where 5677 patients were seen in 2013 with an average 473.08 patients a month and 15.6 a day. Most cases (93%) were dealt with at the health center itself, while 6.8% were referred to the hospital because they were unable to be dealt with properly at the rural health center. Figure 1 shows the chosen health centers in terms of proximity to the regional capital where the hospital is located (ie, Valladolid, Spain).

Although the rural health centers chosen depend on the municipality to whom they provide the service, our choice is based on the premise that they are an acceptable size. The fact that they are centers that lack specialists and/or means available for more complex treatment should also be taken into account. Cardiology cases are referred to the relevant hospital as such cases are not dealt with onsite, as well as other cases of different types.

Therefore, special use of EHRs is made at rural health centers, where the doctor dealing with the consultation modifies the details of the patient being seen. Some patients are then sent on to hospital. The flow of patients at the rural health centers has to taken into account in this study. It is important to note that there are rural health centers with more than one doctor and nurse on duty and even a pediatric and physiotherapy unit. Moreover, the number of patients attended to may vary, reaching up to 400 or 500 a day during doctors' surgery hours, which might be the case with the health centers in Peñafiel and Cuéllar.

Special attention needs to be paid to a swift cloud computing solution for all services, as most of the information being treated will simply be text, meaning that such a flow of data has to be optimized to ensure that electronic consultation and modifications are carried out with as little delay as possible.

Another issue to consider is the fact that the only professional who is authorized to modify EHR information is the doctor, as other medical staff have no access to such information. This is a major factor when consulting the patient's information so as to ensure the privacy of the patients being seen. The doctor is also the only one able to attend to teleconsultation and telediagnoses.

The means and technology exist in these four health centers to implement service providers in each center, along with the infrastructure needed to ensure good quality communication. All four centers have broadband Internet. The parameters for choosing a secure infrastructure on the cloud are the cost per unit of time and online traffic.

XSL•FO
RenderX

Figure 1. Outline of rural health centers close to the hospital in Valladolid.



Methods

iCanCloud [24] software is used to carry out the simulations within the proposed scenario, which provides a type of modeled physical structure that the user can work with. Using this tool significantly reduces the cost of contracting resources because the user can perform a range of simulations of their future infrastructure until they achieve an efficient result in terms of cost/unit of time. We note here that some of the data from each health center are approximate when choosing the infrastructure required to provide access to each eHealth service (eg, regarding the volume of data being handled in each center). This information is required when offering an infrastructure to the EHR system. For telediagnosis and teleconsultation services, the most important criteria when selecting the technology are quality of the image and availability of the system, in real time, as in the case of a telecardiology solution. A secure system has been chosen for all solutions that provides access to the information handled (both text and images), which can be accessed via authentication on the part of the user—both patient and doctor.

Results

Results

Here we describe and compare the key elements associated with eHealth services offered via the cloud and the infrastructures required in terms of security for each. The first stage is common to all services with respect to secure authentication by patients and doctors.

Secure Authentication for All Services

Regardless of the type of eHealth application that needs to be implemented together with the cloud computing technology, all solutions must be correctly authenticated. This is a fundamental, indisputable factor that needs to be covered, as otherwise data might be used by unauthorized persons, which would entail loss of information, fraudulent use, etc, which is unacceptable for this type of an environment. In all cases, a private cloud has been chosen in order to increase security measures. Therefore, diverse encoding technologies, digital signatures, tools, and technologies have been researched to offer an effective, viable solution for correct authentication by health professionals. The decision has also been made to implement the use of smartcards in the system as a feasible and efficient solution for maintaining identity with a suitable degree of security for this joint cloud computing system for any eHealth application introduced on the cloud. Figure 2 shows an example of a consultation in which operation of the doctor's computer and the CAD card reader are observed in detail.

The computer(s) located in the professional health staff's rooms are equipped with a Card Acceptance Device (CAD), which is a smartcard reading device that takes charge of the customer part of the system. This is sufficient for a Java Virtual Machine (JVM) and an OpenCard Framework bookstore (OCF). The OpenCard Framework or OCF of a smartcard is simply middleware implemented in Java that enables an application to be aware of the card's presence and to interact with it in accordance with ISO/IEC standards 7816-4, -8, and -9 [25]. All the aforementioned interact jointly with health care software from the system in each terminal where we have the CAD installed.

OCF is developed by OpenCard Consortium, including IBM and other leading firms in the sector. The framework is implemented in Java and provides an environment that can be developed independently of the card manufacturer. In terms of architecture, OCF is located between the CAD and the host application in the computer. It is hoped that OCF will be

XSL•FO
RenderX

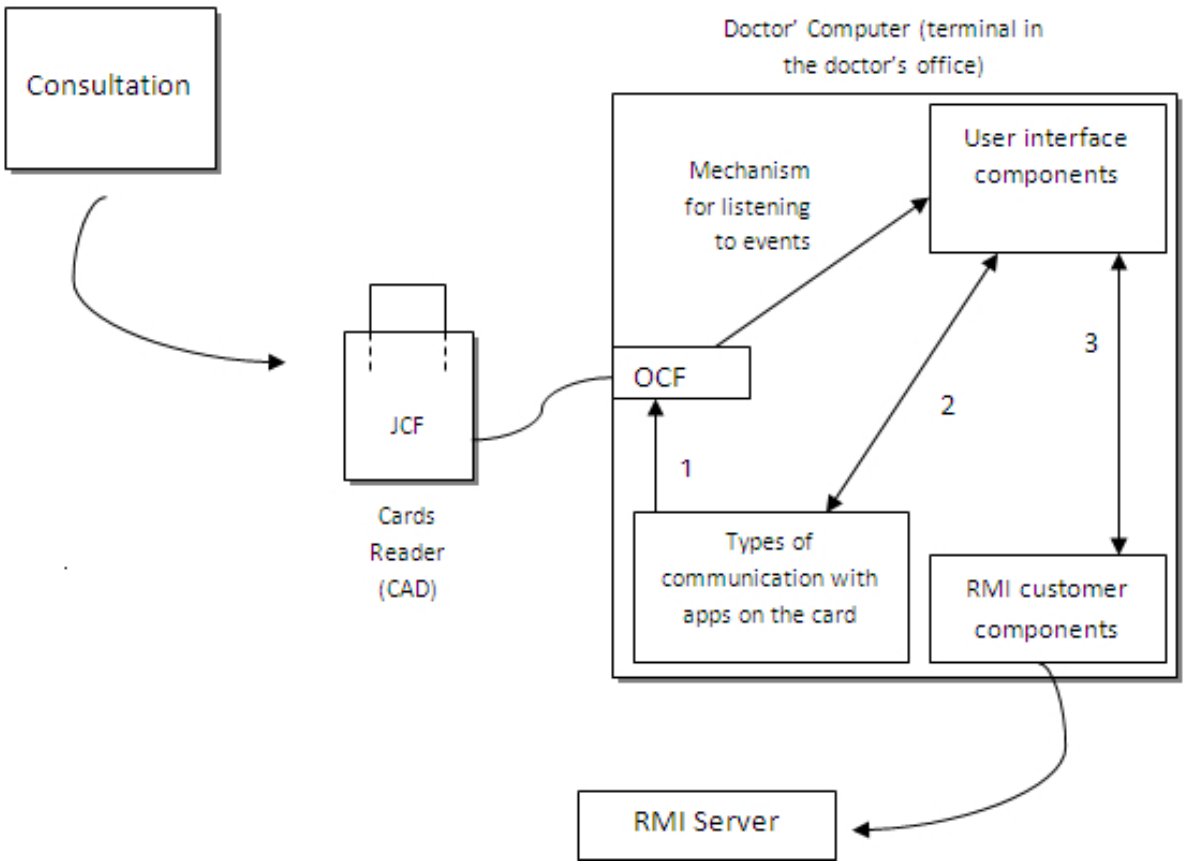
integrated in smartcards destined for use in health care systems [26].

The customer smartcard package that is run in the customer equipment ensures communicates both for the doctor's card and for the patient's. It prepares the environment for card communication with the computer via the Application Protocol Data Unit (APDU) protocol, that is, the ISO 7816 smartcard communication protocol. This package is designed as an API to provide what we might refer to as summaries of customer cards, including APDU communication interfaces. In this way, the components of the user interface are able to use objects provided by this package without having to continually struggle with conversations about smartcard data structure.

An object known as the Card Manager take charges of initializing and shutting down the computer, and establishing a secure communication channel between the smartcard and the session that happens to be open at a specific time. The data regarding sessions are handled by objects known as DoctorSession and PatientSession, which communicate with each other as explained via APDU.

The smartcard terminals act as customers in the remote method invocation (RMI) protocol by calling remote object methods. As previously mentioned, the customer equipment does not contain any software in charge of accessing databases and making enquiries, meaning that the customer software only contains the components of the user interface (eg, types of Java) and ways of displaying layers of Model View Controller (MVC) architecture.

Figure 2. Sample consultation showing operation of smartcards.



Electronic Health Record Service: Infrastructure and Features

The aim is to provide support to an EHR application via cloud computing within the rural health center environment. A series of aspects need to be taken into account prior to considering any solution or implementation. Specifically, the fact that the solution should be secure and robust must be taken into account, and privacy must be provided via efficient authentication support given to ensure that EHRs can be accessed by authorized health care staff. The storage system needs to be scalable yet fair. There is a small number of health records to be dealt with, according to the rural population registered at the center. The data must be replicated, as it is of a sensitive nature, and this process will be repeated throughout the pairs of application scenarios. It is vital that the information be backed up against any possible fraudulent use that may give rise to loss, or in case the information itself becomes mislaid. The quality of communication needs to be of acceptable standard, and there should also be a fiber service to prevent any delays in communication as far as possible. We will likely experience a certain network overload when making simultaneous requests for access to EHRs, meaning that a decent network infrastructure will not manage to avoid everything, although it will do so to a large extent. Each session involving access to data needs to be assured in detail, from access to the session to shutting it down—which must be secure and never exposed to any possible fraudulent use of the information. Although there may be major

XSL•FO
RenderX

information loads in terms of movement, EHRs are mostly written in flat text. Despite the fact that it may sometimes include images from specialists, it is assumed that data traffic will be at an average level. This will affect the structure of the proposed network and, as a result, the storage system.

A general outline of the solution is as follows. The first essential step is access to the cloud by the doctor or professional health care worker. In the case of the solution, access to data can be gained only if the doctor is correctly authenticated in the system, otherwise no secure connection can be established with the information system where the EHRs are located.

The general connection outline within each rural center is shown in Figure 3. Research has been carried out into offering a secure solution using cloud computing technology—an eHealth application fully implemented on the cloud together with the elements required for security and storage—so as to supply the service with EHRs on the cloud for rural health centers. Figure 4 shows the ideal implementation.

The outline of each health center to be connected to the cloud is represented in Figure 5. Each data input and output from and to the cloud passes through two Cisco firewalls with Failover IP configuration in order to obtain a secure and robust service 24/7.

We will subsequently find the input and output router for each rural health center to the cloud. For the solution involving implementation of EHRs, we decided to always install load balancers in communication together with several Internet service providers, due to the fact that we are unaware of the exact amount of information to be received.

Each center will have two Cisco firewalls in Failover IP configuration so as to provide the system with a secure and uninterrupted service. This ensures secure access to the system in the cloud and a shielded communication system between the health centers and the storage system on the cloud. The communication channel is assured via the input/output router of each health center with the cloud. Figure 6 shows the outline for implementation of the EHR system. These elements will have been used previously in other solutions as well as the infrastructure required for each consultation, to enable the smartcard to be used by doctors and the local area network (LAN) infrastructure to be established at the center for the output of data to the outside.

In Figure 7, the infrastructure on the cloud for the full implementation of EHR system is shown. A scalable solution is proposed in terms of capacity for information, which is secure thanks to the shielded communication channels created. Privacy is also granted via use of smartcards exclusively held by authorized health care staff from the center. The cost per month of this solution is around €500.

A LAN connected to the communication output elements will be available at any of the health centers, whereby each doctor may use their smartcard in the CAD via each consultation to obtain the precise information they need on each patient. The problems associated with simultaneous requests from various health centers are eliminated thanks to the load balancers and the use of several Internet service providers who will show the result to the doctor when they require it. The request for information about health records will be dealt with by the database service provider, which understands that there is a staff member authorized to be authenticated using the smartcard system, meaning that the information shown by the Internet service provider can be displayed by the doctor.

XSL•FO
RenderX

Figure 3. Connection outline for each rural center.

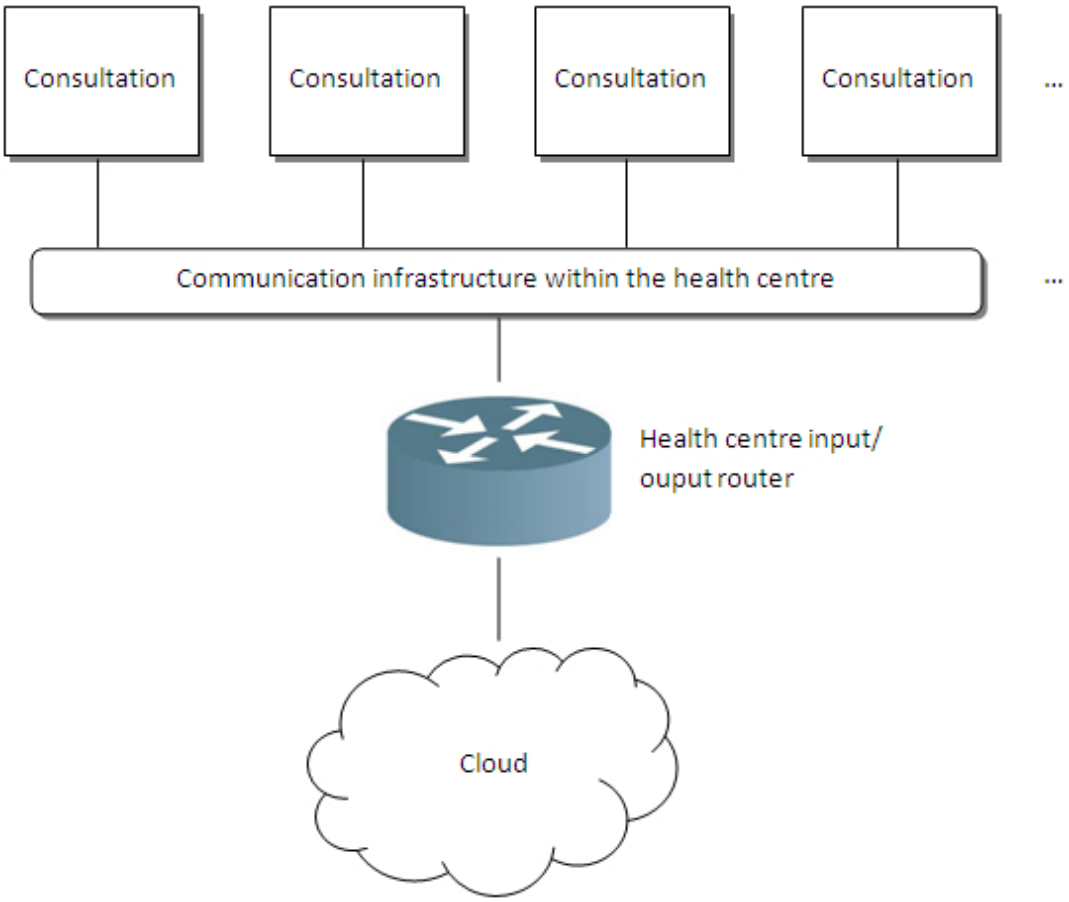
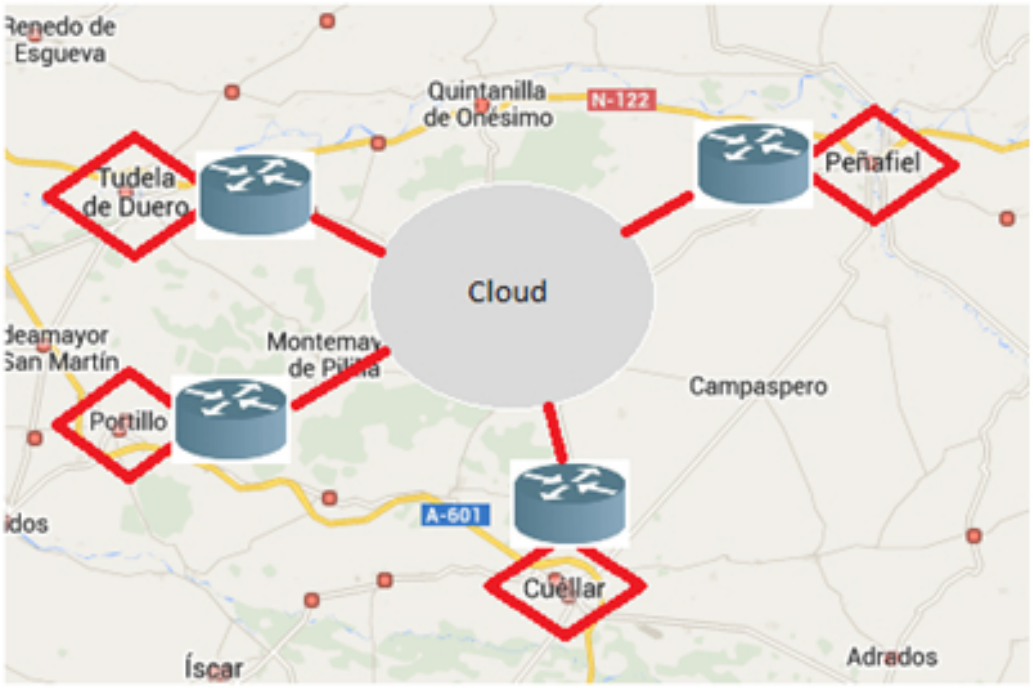
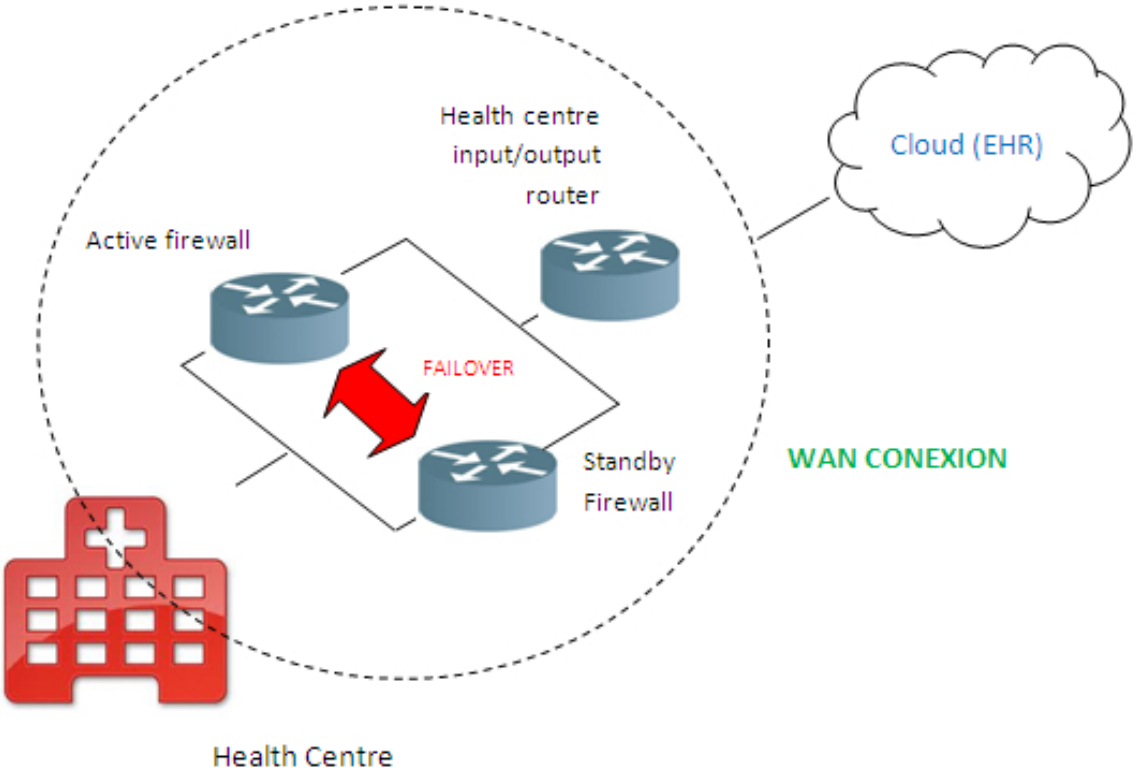


Figure 4. General connection outline for the rural health centers with the cloud.



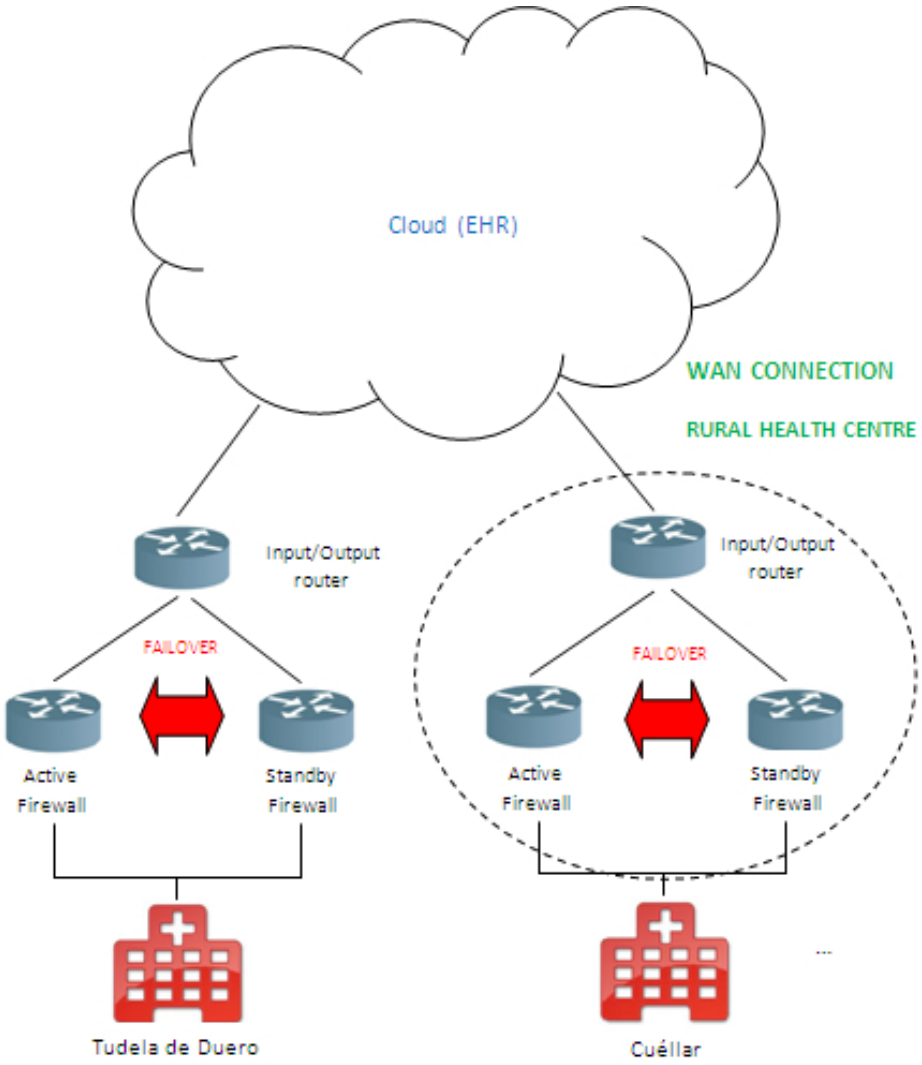
XSL•FO
RenderX

Figure 5. General connection outline for each rural health center.



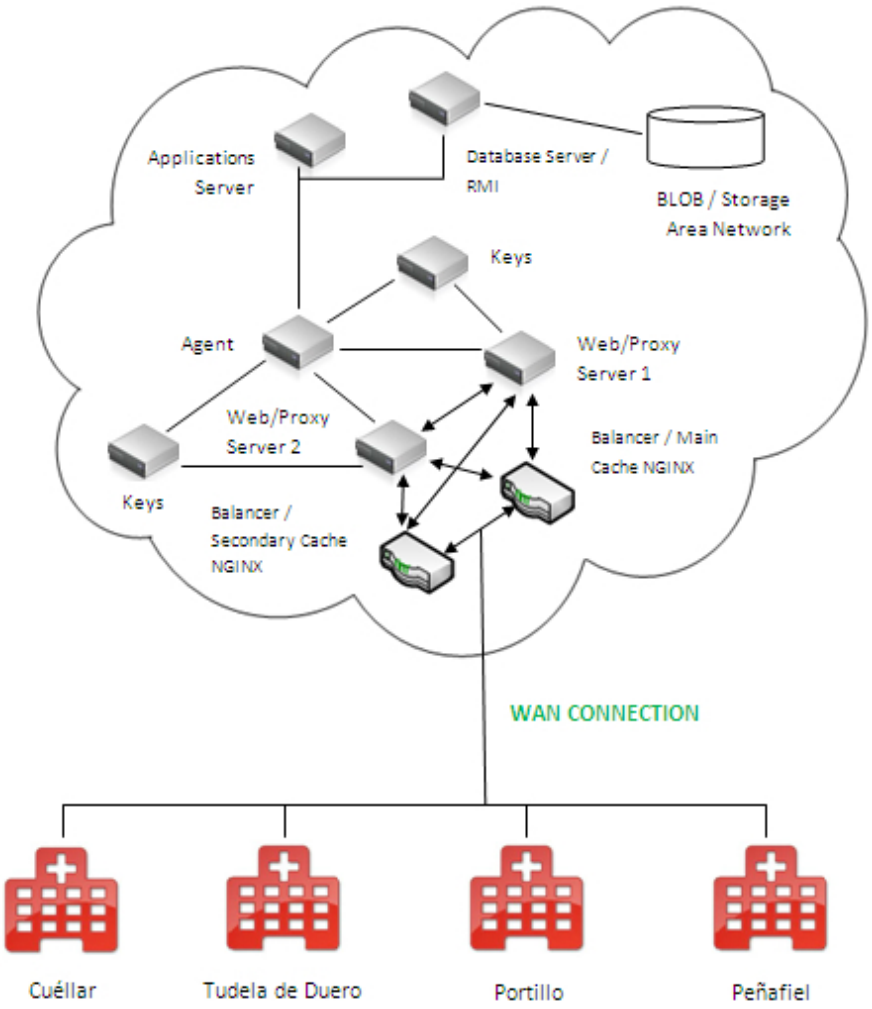
XSL•FO
RenderX

Figure 6. General connection outline for health centers with the cloud.



XSL•FO
RenderX

Figure 7. Outline of the infrastructure on the cloud for full implementation of EHRs.



Telecardiology Service: Infrastructure and Features

The differences between the telecardiology service that has been fully implemented on the cloud in rural health centers and implementation of the previous eHealth application on the cloud is explained in this section.

Those features that are similar to full implementation of EHRs on the cloud are as follows: the wide area network (WAN) connection and rural health center infrastructure, and secure authentication using smartcards. Storage retains the same infrastructure and filing systems, as well as access to them and the StorageGRID technology to administer them. The database and Internet service providers perform the same tasks together with the agent, and the load balancers will remain essential for this solution.

The differences with EHR implementation are as follows: there must be communication with the hospital in order to obtain telecardiology support. As we are dealing with a rural environment, this is essential given that the rural health center does not provide this service. A new LAN infrastructure is needed, if one did not previously exist, to provide connection support for electrocardiograms, echocardiograms, Holters, and angiocardiographs (ie, all the electronic tools required for cardiology diagnosis on the communication side of the hospital). A cardiology connection infrastructure is required at the closest hospital with the cloud so as to provide telecardiology support for the rural health centers.

Load balancers and several Internet service providers will be used to provide support for all the requests entering the system on the cloud, as this could be a large number due to the unevenly distributed population within the municipalities. The filing system and also the security methods are as in the previous models; therefore, the cost per month of such a solution is around €450.

In this scenario, we decided that the health centers should have direct contact with the hospital via the cloud. In this way, doctors from the rural health centers may upload data onto the cloud and the specialist in turn may upload images they as they see fit onto the filing system. They can also leave messages for the doctor or consult the doctor from the rural center for diagnosis of some specific case regarding the patient. We also decided to introduce this system to ensure that the telecardiology application goes directly to the cardiologist at the hospital, so that they will be able to offer better quality treatment to each specific patient with direct care provided from the hospital. Using electronic tools provided by cardiologists from the hospital, the image can then be created locally and subsequently uploaded onto the cloud. Once the application service provider has finished processing the image, it records it in the database and returns the result to the Internet service provider so that the

XSL•FO
RenderX

specialist or general practitioner may display this result and make the relevant decisions in each given case.

The following are used: secure information channels using firewalls and the proxy/Internet-agent with respective coded keys, the BLOB + SAN (RAID 0+1) filing system with back-up copy administered using StorageGRID, which is able to make back-up copies in different physical devices, and smartcards for authentication of professional health staff. In this way, a fully shielded solution is obtained for using patient data both on the communication side of the health center and by the specialist at the hospital (if required).

Regardless of where the health center is located, the cloud is always available for attending to a cardiology service from the closest hospital. Thanks to the BLOB, the data service provider will be unburdened in attending to requests from different centers without overload problems. Similarly, the application/transaction service provider provides an optimum flow of data because it uses a message queuing system to create a solid flow that it is able to process.

Teleconsultation and Telediagnosis Service: Infrastructure and Features

Teleconsultation and telediagnosis are services with a major flow of traffic—an important factor when seeking a suitable cloud computing solution. The infrastructure itself on the cloud must be prepared to support a major information traffic load, as we start from the premise that the hospital infrastructure is prepared to support major loads of this type as in previous cases.

Doctors may consult specialists and specialists may consult other specialists or there may be a variety of combinations that require a tolerant system with a properly structured major work load to be implemented.

These are the two costliest eHealth services due to the fact that they have a major influence over the rest. This leads to the search for a less costly solution, whereby a sufficient cloud computing scenario is proposed that can at the same time be adapted as much as possible without including an excessive number of elements that could be detrimental by giving rise to more traffic, as well as raising the cost of the infrastructure.

The solution adapts perfectly to the previous case in which the specialist could be consulted and images uploaded by the cardiologist, etc. We consider the same infrastructure as in the previous case since consultation and diagnosis may be required of specialists by the hospital. This decision was made for different reasons, as load balancers and Internet service providers would be needed for a possibly major flow of traffic from different rural health centers. Diagnosis or consultation by specialists located at the hospital may prove to be necessary. Therefore, we decided to implement the same infrastructure as in the previous case so as to be able to maintain communication with the hospital and for specialists to upload images, send messages to doctors, respond to inquiries, and help with diagnoses, etc. This will help improve the quality of the patient's treatment.

The doctor may access the patient's data in the same way as in the previous case, by modifying, displaying, or updating information about the latter via the use of smartcards by authorized professional health care staff—in this specific case, the doctors at the rural health center.

The solution to this scenario is adapted to the requirements offered by teleconsultation or telediagnosis. As the information is always made available on the cloud via online portals, there is open communication with specialists and support can be provided regardless of location. This translates into a major cost in infrastructure for the health care body and an improvement in the patient's quality of life.

Now that each of the secure solutions for the different eHealth services has been described, Table 1 shows a comparison between the three solutions according to criteria such as type of service, type of cloud, estimated cost per unit of time, estimated amount of traffic on the network, and key security elements.

Table 1. Comparison between cloud-based solutions for the different eHealth services.

Type of service	Type of cloud	Estimated cost (€/month)	Traffic on the network (Gbps)	Key security elements
EHRs	Private	500	1	Firewalls
				Load balancers
				Smartcards
				Filing system
Telecardiology	Private	450	1	Firewalls
				Load balancers
				Smartcards
Teleconsultation/ Telediagnosis	Private	500	2	Firewalls
				Load balancers
				Smartcards

XSL•FO
RenderX

Discussion

Principal Considerations

After researching cloud computing technology to discover whether secure solutions can be offered to telecardiology, teleconsultation, telediagnosis, and EHR eHealth services on the cloud for different rural health centers, we have offered an optimum infrastructure as a viable proposal for each given case. Access to these services on the cloud enables more reliable treatment and diagnosis to be offered, above all in environments that lack certain services such as health centers in small municipalities where no medical specialists exist.

We opted for a private cloud for all services in order to further ensure security levels. The costs of the infrastructure network at the four health centers are estimated, with the costliest being those assigned to the EHR and teleconsultation and telediagnosis service. As far as the estimated amount of network traffic is concerned, the service that gives rise to the most traffic is that of diagnosis, owing to the images and videos sent via the network.

The use of computing technology on the cloud alongside the proposed smartcard system for doctors' authentication helps health centers to electronically manage all health-related data about patients and enables them to make any modifications in a reliable manner—meaning that privacy, security, and robustness are assured in an extremely sensitive data system.

Any authorized staff member, doctor, or other health care professional may access the services provided by the different eHealth applications at any time and from any location within the different scenarios proposed, under the assurance that their privileges be maintained on the cloud by using their smartcard to access data.

Control of access to the system, and the use of Cisco firewalls configured using the Failover IP, greatly enhance security offered by the use of coded keys via proxy and the agent. The storage system proposed enables back-up copies to be created on physically independent elements on the cloud. Together, they increase the privacy and security of all communications from start to finish, as well as ensuring the robustness of the data.

Conclusions

The research carried out for this paper on cloud computing technology, in addition to other technologies required to provide excellent authentication of the system, has allowed us to suggest a solution adapted to each eHealth service at rural health centers that offers security, privacy, and robustness and can also be deemed optimum for a large number of requests on the cloud.

Within a common scenario of cloud configuration, a customer is initially unaware of the requirements they need to provide infrastructure to the software housed, and even less so at any optimum level. One of the future areas of research would be to analyze the eHealth services proposed in this paper in other scenarios such as hospitals or an individual health center.

Acknowledgments

This research has been partially supported by Ministerio de Economía y Competitividad, Spain.

Conflicts of Interest

None declared.

References

1. National Institute of Standards and Technology. Computer Security Resource Center. 2015. URL: http://www.csrc.nist.gov(last [accessed 2015-03-19] [WebCite Cache ID 6ZP04lLKf]
2. Cloud Security Alliance. 2009. Security Guidance for Critical Areas of Focus in Cloud Computing, April URL: http://cloudsecurityalliance.org/guidance/csaguide.v1.0.pdf(last [accessed 2015-03-19] [WebCite Cache ID 6ZP073259]
3. Cloud Computing, CSA (Cloud Security Alliance). Cloud Computing CSA URL: http://www.cloudsecurityalliance.org(last [accessed 2015-06-19] [WebCite Cache ID 6ZP0D3WXL]
4. Silva BM, Rodrigues JJPC, Canelo F, Lopes IC, Zhou L. A data encryption solution for mobile health apps in cooperation environments. J Med Internet Res 2013;15(4):e66 [FREE Full text] [doi: 10.2196/jmir.2498] [Medline: 23624056]
5. Kim JT. Enhanced secure authentication for mobile RFID healthcare system in wireless sensor networks. Communications in Computer and Information Science 2012;352:190-197.
6. Marston S, Li Z, Bandyopadhyay S, Zhang J, Ghalsasi A. Cloud computing: the business perspective. Decis Support Syst 2011;51(1):176-189.
7. Guo L, Zhang C, Sun J, Fang Y. A Privacy-Preserving Attribute-based Authentication System for Mobile Health Networks. IEEE Trans Mob Comput 2014;13(9):1927-1941.
8. Luxton DD, Kayl RA, Mishkind MC. mHealth data security: the need for HIPAA-compliant standardization. Telemed J E Health 2012 May;18(4):284-288. [doi: 10.1089/tmj.2011.0180] [Medline: 22400974]
9. Coles-Kemp L, Reddington J, Williams P. Looking at clouds from both sides: The advantages and disadvantages of placing personal narratives in the cloud. Information Security Technical Report 2011;16:115-122.
10. Kuo AM. Opportunities and challenges of cloud computing to improve health care services. J Med Internet Res 2011;13(3):e67 [FREE Full text] [doi: 10.2196/jmir.1867] [Medline: 21937354]

XSL•FO
RenderX

11. Marks E, Lozano B. Executive's Guide to Cloud Computing. Hoboken, NJ: Wiley; 2010.
12. Sadeghi AR. Mobile security and privacy: The quest for the mighty access control. 2013 Presented at: ACM Symposium on Access Control Models and Technologies, SACMAT; 2013; Amsterdam p. 1-2.
13. Piette JD, Mendoza-Avelares MO, Ganser M, Mohamed M, Marinec N, Krishnan S. A preliminary study of a cloud-computing model for chronic illness self-care support in an underdeveloped country. Am J Prev Med 2011 Jun;40(6):629-632 [FREE Full text] [doi: 10.1016/j.amepre.2011.02.014] [Medline: 21565655]
14. Low C, Hsueh CY. Criteria for the evaluation of a cloud-based hospital information system outsourcing provider. J Med Syst 2012 Dec;36(6):3543-3553. [doi: 10.1007/s10916-012-9829-z] [Medline: 22366976]
15. Mather T, Kumaraswamy S, Latif S. Cloud Security and Privacy: An Enterprise Perspective on Risks and Compliance (Theory in Practice). Sebastopol, CA: O'Reilly Media, Inc; 2009.
16. Wei J, Hu X, Liu W. An improved authentication scheme for telecare medicine information systems. J Med Syst 2012 Dec;36(6):3597-3604. [doi: 10.1007/s10916-012-9835-1] [Medline: 22374237]
17. Tejero A, de la Torre I. Advances and current state of the security and privacy in electronic health records: survey from a social perspective. J Med Syst 2012 Oct;36(5):3019-3027. [doi: 10.1007/s10916-011-9779-x] [Medline: 21947856]
18. Fernández-Cardeñosa G, de la Torre-Díez I, López-Coronado M, Rodrigues JJPC. Analysis of cloud-based solutions on EHRs systems in different scenarios. J Med Syst 2012 Dec;36(6):3777-3782. [doi: 10.1007/s10916-012-9850-2] [Medline: 22492177]
19. Rodrigues Joel J P JJPC, de la Torre I, Fernández G, López-Coronado M. Analysis of the security and privacy requirements of cloud-based electronic health records systems. J Med Internet Res 2013;15(8):e186 [FREE Full text] [doi: 10.2196/jmir.2494] [Medline: 23965254]
20. Cloud Computing: Top 5 Security Concerns. Health Information Technology, Implementation, Insight, News, Spotlight, Today. URL: http://www.ehrscope.com/blog/cloud-computing-top-5-security-concerns(last [accessed 2015-06-19] [WebCite Cache ID 6ZP0hB7oe]
21. Zhang R, Liu L. Security Model and Requirements for Healthcare Application Clouds. In: Proceedings of the IEEE 3rd International Conference on Cloud Computing. 2010 Presented at: IEEE 3rd International Conference on Cloud Computing; July; 2010; Miami, Florida.
22. Church P, Goscinski A, Lefèvre C. Exposing HPC and sequential applications as services through the development and deployment of a SaaS cloud. Future Gener Comp Sy 2015;43:44-37.
23. Anselmi J, Ardagna D, Passacantando M. Generalized Nash equilibria for SaaS/PaaS Clouds. Eur J Oper Res 2014;236(1):326-339.
24. Núñez A, Vázquez-Poletti J, Caminero A, Castañé G, Carretero J, Llorente I. iCanCloud: A Flexible and Scalable Cloud Infrastructure Simulator. J Grid Comput 2012;10(1):185-209.
25. OpenCard Framework (OCF). URL: http://www.openscdp.org/ocf(last [accessed 2015-06-19] [WebCite Cache ID 6ZP0jnfDe]
26. Rankl W, Effing W. Smartcard Handbook, second ed. West Sussex, England: John Wiley & Sons; 2000:5-50.

Abbreviations

APDU: Application Protocol Data Unit
CAD: Card Acceptance Device
CSA: Cloud Security Alliance
DDoS: Distributed Denial of Service
EHR: electronic health record
IaaS: Infrastructure as a Service
ICT: Information and Communications Technologies
JVM: Java Virtual Machine
LAN: Local Area Network
MVC: Model View Controller
OCF: Open Card Framework
PaaS: Platform as a Service
QoS: Quality of Service
SaaS: Software as a Service
SSL: Secure Sockets Layer
WAN: Wide Area Network

XSL•FO
RenderX

Edited by G Eysenbach; submitted 12.03.15; peer-reviewed by K Lin, J Herreros; comments to author 01.04.15; revised version received 13.04.15; accepted 17.04.15; published 27.07.15.

Please cite as:
de la Torre- Díez I, Lopez-Coronado M, Garcia-Zapirain Soto B, Mendez-Zorrilla A
Secure Cloud-Based Solutions for Different eHealth Services in Spanish Rural Health Centers
J Med Internet Res 2015;17(7):e157
URL: http://www.jmir.org/2015/7/e157/
doi:10.2196/jmir.4422
PMID:26215155

©Isabel de la Torre- Díez, Miguel Lopez-Coronado, Begonya Garcia-Zapirain Soto, Amaia Mendez-Zorrilla. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 27.07.2015. 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 Cost-Effectiveness Analysis of Blended Versus Face-to-Face Delivery of Evidence-Based Medicine to Medical Students

Stephen Maloney[1], PhD; Peter Nicklen[1], B.Physio; George Rivers[2], PhD; Jonathan Foo[1]; Ying Ying Ooi[1], M.Comm, CPA; Scott Reeves[3], PhD; Kieran Walsh[4]; Dragan Ilic[5], PhD

[1]Monash University, Frankston, Australia

[2]Monash University, Caulfield, Australia

[3]Centre for Health & Social Care Research, Kingston University & St George's, University of London, London, United Kingdom

[4]BMJ Learning, London, United Kingdom

[5]Monash University, Melbourne, Australia

Corresponding Author:
Stephen Maloney, PhD
Monash University
PO Box 527
Frankston,
Australia
Phone: 61 3 990 44240
Fax: 61 3 990 44812
Email: stephen.maloney@monash.edu

Abstract

Background: Blended learning describes a combination of teaching methods, often utilizing digital technologies. Research suggests that learner outcomes can be improved through some blended learning formats. However, the cost-effectiveness of delivering blended learning is unclear.

Objective: This study aimed to determine the cost-effectiveness of a face-to-face learning and blended learning approach for evidence-based medicine training within a medical program.

Methods: The economic evaluation was conducted as part of a randomized controlled trial (RCT) comparing the evidence-based medicine (EBM) competency of medical students who participated in two different modes of education delivery. In the traditional face-to-face method, students received ten 2-hour classes. In the blended learning approach, students received the same total face-to-face hours but with different activities and additional online and mobile learning. Online activities utilized YouTube and a library guide indexing electronic databases, guides, and books. Mobile learning involved self-directed interactions with patients in their regular clinical placements. The attribution and differentiation of costs between the interventions within the RCT was measured in conjunction with measured outcomes of effectiveness. An incremental cost-effectiveness ratio was calculated comparing the ongoing operation costs of each method with the level of EBM proficiency achieved. Present value analysis was used to calculate the break-even point considering the transition cost and the difference in ongoing operation cost.

Results: The incremental cost-effectiveness ratio indicated that it costs 24% less to educate a student to the same level of EBM competency via the blended learning approach used in the study, when excluding transition costs. The sunk cost of approximately AUD $40,000 to transition to the blended model exceeds any savings from using the approach within the first year of its implementation; however, a break-even point is achieved within its third iteration and relative savings in the subsequent years. The sensitivity analysis indicates that approaches with higher transition costs, or staffing requirements over that of a traditional method, are likely to result in negative value propositions.

Conclusions: Under the study conditions, a blended learning approach was more cost-effective to operate and resulted in improved value for the institution after the third year iteration, when compared to the traditional face-to-face model. The wider applicability of the findings are dependent on the type of blended learning utilized, staffing expertise, and educational context.

(J Med Internet Res 2015;17(7):e182) doi:10.2196/jmir.4346

KEYWORDS

evidence-based medicine; economic evaluation; eLearning; medical education

XSL•FO
RenderX

Introduction

Evidence-based medicine (EBM) combines the best available evidence with clinical expertise and patient values [1] and is core to many medical programs worldwide [2-5]. There is an imperative to increase the number of competent EBM trained practitioners. Medical practitioners who are competent and confident in applying EBM possess a powerful tool to inform their decision making. EBM competencies also provide the ability to facilitate life-long learning, as clinicians are able to ask effective clinical questions, acquire information through emerging research, appraise its quality and relevance, apply evidence to practice, and assess its impact [1].

Teaching and learning is undergoing a cultural change with the expansion of student-centered learning, including the push for flipped teaching, peer-assisted learning, increased use of virtual learning environments, and the use of simulation. Research has shown that Web-based learning results in improved outcomes when applied to health professional education [6], with studies focusing on clinical disciplines within medicine reporting an increase in student self-efficacy, knowledge, and self-directed learning [4,7-9]. However, due to the variety of online learning applications and education provider contexts, there is an added need for detailed and robust studies [10], as uninformed transition of learning material to an online environment can have a negative impact on educational outcomes [11]. Online education, or its many variations (eg, eLearning, Web 2.0), appears to hold great promise for addressing the accessibility and efficiency of education, yet currently there is a lack of evidence to inform educators and learners as to the most effective methods of teaching EBM to medical students.

An investigation into the effectiveness of implementing a blended learning (BL) versus a traditional face-to-face (F2F) learning approach of teaching EBM to medical students was conducted by Ilic et al [12]. This multicenter international study used validated outcome measures of EBM competency to determine that BL is no more effective than F2F at increasing medical students' knowledge and skills in EBM. These authors concluded that the BL approach was significantly more effective at increasing student attitudes toward EBM and self-reported use of EBM in clinical practice. Although their study looked at competency, attitudes, skills, and behavior, the missing piece of the puzzle, as acknowledged by the authors, was a measure of cost-effectiveness to help facilitate sustainable adoption of the approach.

The cost and value of teaching and learning practices in medical education have a direct impact on the accessibility of education, the efficiency and quality of education, and the productivity of our health workforce [13-15]. Cost-effectiveness analysis allows decision makers to decrease the risks of implementation from both a financial perspective and the perspective of maintaining quality of education, thereby facilitating adoption [16]. Considerations of cost and value are most commonly known for their application to pharmaceutical interventions: impact of drug A versus drug B, with consideration of the drug's effectiveness, side effects, and value for money. Pharmacists can still stock the most expensive choices, and consumers can still purchase them; however, these are informed decisions based on the evidence. The same principles should apply to education. For an educator to effectively review their practices and pedagogy, they must consider the learning experience and learning outcomes alongside measures of cost and value [17].

Current literature on the cost and value of technological innovations in education is divided, with some studies supporting its cost-effectiveness [18], and others indicating that costs are higher as a result of increased resource development time and need for technological support [19,20]. Economic analysis for health education has primarily focused on telemedicine technology, medical reviews by remote physicians [21], or has been concerned with the cost-effectiveness of modalities for patient education [22]. Previous cost studies on BL approaches in health professional education have typically been conducted on short courses with small sample sizes [18,19] or were unable to determine the cost-effectiveness relationship [23,24]. Thus, there is considerable doubt as to whether existing literature could be generalized to the core-teaching content of a contemporary medical program.

This paper presents the findings from a study that aimed to compare the cost-effectiveness of a face-to-face approach and blended learning approach for EBM training within a medical program—training that has rigorously evaluated effectiveness as tested within a randomized controlled trial (RCT).

Methods

Design

The economic evaluation was conducted as part of an RCT comparing the EBM competency of medical students who participated in two different modes of education delivery [12]. The attribution and differentiation of costs between the interventions within the RCT, in conjunction with measured outcomes of effectiveness, enabled a cost-effectiveness analysis to be applied.

Trial Participants, Methods, and Results

The fundamental elements of the RCT have been reported to provide context for the application of the cost-effectiveness analysis. Further detail on the methods and results of the RCT are available within a published paper [12] and pilot study [4].

A multi-campus study was performed with medical students enrolled in the Bachelor of Medicine, Bachelor of Surgery course at Monash University, Melbourne and Malaysia, including both undergraduate and post-graduate students. Participants were third-year medical students, who were all entering their first year of clinically based training and first year of formal EBM training. The EBM unit is integrated within the medical curriculum and is equivalent to a 6-credit point unit. A total of 497 students were randomized to receive EBM teaching via either the incumbent face-to-face approach (F2F) or the blended learning approach (BL). Students randomized to the intervention group received the same theoretical concepts taught in the control group, but in a BL approach. The BL approach to teaching EBM integrated (1) classroom activities (lecture/tutorial) with (2) online and (3) mobile learning as described in Table 1.

XSL•FO
RenderX

Table 1. Detailed description of the DL and BL teaching approach to EBM with associated costs.

Teaching	Description	Contribution to costs
F2F classroom activities	Student receive 10 x 2 hour classes. Each class starts with the tutor presenting on the EBM content for the session. Students then complete small group tasks and participate in large group discussions lead by the tutor.	F2F classroom activities contribute to staff preparation and teaching time, as well as space charges.
BL classroom activities	At the start of semester, students receive a 4-hour workshop on EBM concepts and an introduction to the BL format. The remaining classes are run in small group format with topics for discussion set by the tutor. Tutors facilitate peer-to-peer learning with a quasi-journal club delivery method where students are assigned topics to investigate during the week. Students then report on their findings in the next session.	BL classroom activities contribute to staffing preparation and teaching time, as well as space charges. Staff have a less active role in the activities, acting as facilitators than tutors. Thus, compared to DL classroom activities, the preparation and teaching costs are lower.
BL online activities	Online activities include a YouTube channel and an online Monash University library guide. The YouTube channel has 11 online lectures with an average length of 17 minutes [25]. The library guide indexes online resources (eg, databases, textbooks, and guidelines) and instructs students on how to use them [26]. Students are sent online activities to complete prior to the classroom session and can also use the online resources to assist learning during the week.	The YouTube channel was developed by Monash University staff solely for the teaching of the EBM unit, making up a large portion of the transition costs. Library staff generate online guides for many subjects. The guide used for EBM teaching existed prior to the BL transition. Thus, costs were attributed to designing activities using the online guide resources, but not the creation of the online guide.
BL mobile learning	Mobile learning occurs on the wards, where students interact with patients during their existing day-to-day "beside teaching" schedule—a method previously piloted [4]. Based on their assigned topic for the week, students are required to identify a patient, take a detailed history, and apply the principles of EBM relevant to the patient.	Mobile learning is completely student self-directed. There were no costs associated with mobile learning factored into the cost model.

A total of 147 (29.6%) of the 497 students completed the follow-up assessments on EBM competency and attitudes. EBM competencies were assessed using the validated Berlin questionnaire [27]. Students' self-efficacy, attitudes, and behavior were also assessed. EBM competency did not differ significantly between students receiving the BL approach versus those receiving the F2F approach: mean difference -0.68, 95% CI -1.71 to 0.34, *P*=.19. Although student ratings of self-efficacy, attitudes, and behaviors all displayed a significant preference for the BL model. In total, 74 students completed the F2F model of training, with a mean score of 7.98 (SD 3.35), and 73 students completed the training via the BL approach, with a mean score of 8.67 (SD 2.96).

Economic Analysis Procedure

The following analysis was applied from the perspective of Monash University, measuring the cost of training student clinicians against their self-reported level of EBM competence. The primary outcome was student competency in EBM, measured 1 month after the teaching activities, using the validated Berlin Questionnaire. We calculated the cost-effectiveness for each course delivery method by first determining the quality of students' education with each method, known as quality-adjusted students educated (QASE), using the formula QASE = number of students educated x the group's average rating on the Berlin Questionnaire. In this approach, the reported average rating was used as a surrogate for measuring the improved ability of the total cohort of 497 students for each teaching approach. QASE is the measurement of effect in the incremental cost-effectiveness analysis. Cost-effectiveness was calculated using the incremental cost-effectiveness ratio (ICER), which measures cost per QASE (Figure 1). The ICER is reflective of the ongoing operational costs and does not include the initial transition cost in its calculation. Thus, the results represent the cost-effectiveness of the second iteration and onwards only.

Figure 1. Equation for the calculation of the incremental cost for each quality-adjusted student educated (ICER).

$$\frac{\Delta S}{\Delta QASE} = \frac{(Operation\ Costs_{intervention} - Operation\ Costs_{control})}{(QASE_{intervention} - QASE_{control})}$$

Two rounds of ICER calculations were completed. First, an ICER comparing F2F to BL was calculated, establishing the hierarchy of cost-effectiveness. There were four possible outcomes of this analysis: (1) BL is more costly and more effective than F2F, (2) BL is more costly and less effective than F2F, (3) BL is less costly and more effective than F2F, and (4) BL is less costly and less effective than F2F [28].

The World Health Organization recommends that cost-effectiveness comparisons should be carried out against a common baseline as this is more comparable across populations and studies [29]. Thus, in the second round of ICER calculations, F2F and BL were independently compared to no EBM training. As there are no pre-test Berlin scores from the RCT, baseline scores from third-year medical students with no prior EBM training (mean 4.2, SD 2.2) from the Berlin Questionnaire validation study by Fritsche et al were used [27]. Independent ICER calculation of F2F and BL compared to baseline calculates the cost per student per increase in QASE. Due to commercial sensitivities, these values were compared and the percentage difference reported.

The sunk cost of transitioning to a BL format was not included within the ICER, as is typical within economic analyses. However, the transition costs are reported separately, due to its importance to decision makers considering implementation or

XSL•FO
RenderX

adoption of similar pedagogy. A further present value (PV) break-even analysis incorporating transition costs was calculated using a real discount rate of 4% (Figure 2). Break-even analysis calculates the point in time at which the total running cost of the F2F approach equals the total running cost of the BL approach plus the cost of transitioning to BL. PV accounts for the time preference of money, allowing for present day comparisons to be made on future cash flows. Subsequent iterations of the program are assumed to occur at 1-year intervals. Due to commercial sensitivities, PV values are expressed as the difference between teaching methods.

Figure 2. Equation for the PV break-even calculation where C=cost of teaching method, r=discount rate, t=number of years, and BL_0=cost of transitioning to BL.

$$\sum_{t=1}^{T} \frac{C_{F2F}}{(1+r)^t} = \sum_{t=1}^{T} \frac{C_{BL}}{(1+r)^t} + BL_0$$

Measurement of Cost

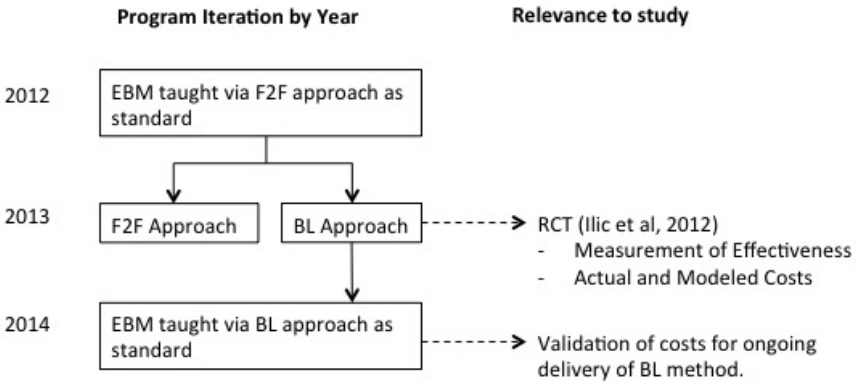
The BL arm of the RCT has continued as the preferred method of delivery of EBM with the medical program. Costs were modeled on the 2013 RCT, and the assumptions relating to the ongoing costs were validated within the 2014 iteration of the program. The course iterations and outputs as they relate to the analysis are shown in Figure 3.

Cost categories considered within the analysis included (1) transition costs, (2) ongoing staffing costs, (3) space charges, and (4) overheads (Table 2). All staffing hours were calculated from actual values observed in the original RCT and differentiated by academic level. Other values where modeled have been explicitly stated in Table 2, along with the assumptions made. Where appropriate, results have been expressed as the difference between F2F and BL due to commercial sensitivities.

Table 2. Description of cost categories.

Cost	F2F	BL	Description
Transition costs	No	Yes	Staff time for creating YouTube resources, tailoring content for BL approach, and designing activities.
Ongoing staff costs	Yes	Yes	Staff time for regular preparation for incoming students, teaching, unit-coordinating, and marking assessments.
Staff on-costs	Yes	Yes	Staffing salaries were assumed to be the highest step within the academic level as per the Monash University Enterprise Agreements [30]. Further on-costs include superannuation contributions, payroll taxes, loading, and levies. Academic staff (professors and associate professors) receive 17% superannuation contribution, with total on-costs of 39.62%. Casual staff receive 9.25% superannuation, hence total on-costs for casual staff is 15.43%.
Space costs	Yes	Yes	Space costs for classroom activities are dependent on seating capacity and AV capabilities, based on Monash University rate charges to faculties. Seating requirements were known from the RCT; however, AV capabilities were modeled as requiring full AV and digital projection. Number of classroom hours booked for each method was obtained from the university online timetable system records.
Overhead costs	Yes	Yes	The total operating overheads were modeled at 37% in line with Monash University's Project Costing and Price Model. Central overheads were calculated at 22% (allocation of corporate services costs: finance, HR, IT, & corporate services). Faculty overheads were valued at 10% for the allocation of professional staff involved in the general support of students including student services, IT support for students, and research activities. Other overheads included 5% allocated to the general costs of running support activities.

Figure 3. Program iterations and outputs relevant to past and current research.



Sensitivity Analysis

We included a multivariate sensitivity analysis around permutations to the key variables. Guided by available literature, scenarios were constructed around increased transition costs and increased staffing requirements for running the BL format. Thus, the robustness of our economic model was tested, allowing the reader to tailor the findings to different educational settings.

XSL•FO
RenderX

Results

BL and F2F Approach Delivery Inputs

The staffing profile, wage, and difference in hours attributed to the delivery of each method are detailed in Table 3 below. Where there is a difference in hours, it is expressed as F2F hours minus BL hours; that is, positive values indicate fewer BL hours.

Space charges for the F2F and BL teaching methods were identical. Both methods had booked rooms seating 180 people (AUD $131.34/hr) for 28 hours and rooms seating 50 people (AUD $76.61/hr) for 210 hours. Major updating of the curriculum content and materials was considered required at 5 years and estimated at 160 hours of staff time for both F2F and BL approaches.

The cost of transitioning from the F2F to BL approach is the result of staffing time plus on-costs, valued at AUD $38,186. The difference in delivery costs is AUD $12,514 in favor of the BL method. The lower BL delivery cost is the result of lower preparation and direct teaching costs compared to the F2F method.

Table 3. Staffing profile with difference in hours for F2F and BL approaches.

Cost	Academic level	Rate, AUD	Difference (F2F hrs - BL hrs)
Transition costs (creating YouTube videos, tailoring content, and activities)	Professor	$84.23/hr	-40 hrs
	Associate Professor	$72.04/hr	-120 hrs
	Casual staff with PhD	$47.57/hr	-100 hrs
	Casual staff no PhD	$39.78/hr	-160 hrs
Total hour difference for transition costs			-420 hrs
Preparation costs (getting ready for each class and regular minor updating)	Professor	$84.23/hr	5 hrs
	Associate Professor	$72.04/hr	15 hrs
	Casual staff with PhD	$47.57/hr	12.5 hrs
	Casual staff no PhD	$39.78/hr	0 hrs
Direct teaching (classroom activities)	Professor	$84.23/hr	5 hrs
	Associate Professor	$72.04/hr	15 hrs
	Casual staff with PhD	$47.57/hr	12.5 hrs
	Casual staff no PhD	$39.78/hr	20 hrs
Unit coordination	Associate Professor	$72.04/hr	0 hrs
Marking and examining	Professor	$84.23/hr	0 hrs
	Associate Professor	$72.04/hr	0 hrs
	Casual staff with PhD	$47.57/hr	0 hrs
	Casual staff no PhD	$39.78/hr	0 hrs
Total hour difference for ongoing costs			85 hrs

Cost Effectiveness

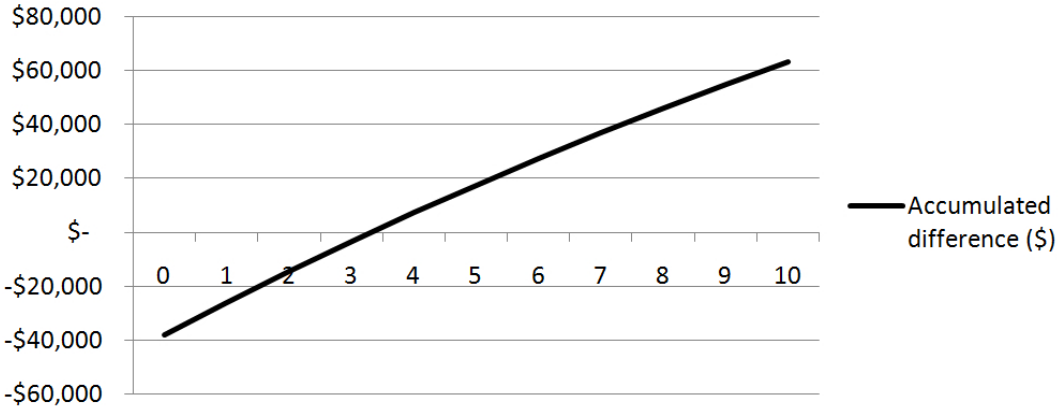
The BL method was less costly and more effective to operate than the F2F approach. The ICER result comparing F2F to BL was -$1.10, indicating that to operate the BL model, there is a saving of $1.10 per student per increase in QASE above the QASE of the F2F method. Independent ICER calculations of F2F and BL compared to baseline found that BL was 24% more cost-effective to operate than F2F. That is, the cost of achieving a statistically similar Berlin score for the same number of students is 24% less in the BL method, excluding transition cost. Given that the BL approach is slightly more effective than the F2F approach, the BL approach would have to cost 18% more than the F2F approach to run for there to be no difference in ICER between the two approaches.

Break-Even Analysis

The PV difference in cost between F2F and BL approaches, including the transition costs to the BL model, are shown in Figure 4. The graphed line demonstrates the sunk cost of the transition to the blended learning approach, the cost of which is recovered by the lower running cost of the BL method in the third year. Subsequent years show a relative saving using the BL method, with PV savings of approximately AUD $17,000 after 5 years and AUD $63,000 after 10 years.

XSL•FO
RenderX

Figure 4. The accumulated difference in the PV of cost between F2F and BL across 10 years.



Sensitivity Analysis

A multivariate sensitivity analysis was conducted with scenario variations to ongoing staffing costs and mean Berlin score (Table 4). Results are reported as the percentage difference in the independent F2F and BL ICER values. Two scenarios of educational effectiveness with the BL approach were simulated, that of mean Berlin scores as per the RCT and another with scores equivalent to the F2F method. The results show the BL format is more cost-effective to operate up to staffing levels 20% higher than F2F, using the RCT mean Berlin score.

Table 4. Multivariate sensitivity analysis adjusting staffing levels and mean BL Berlin score.

		ICER % difference by mean BL Berlin score[a]	
Scenario variation	Adjustment made	Effectiveness as per RCT	Equivalent effectiveness
Current model	Nil	24%	10%
Same preparation time	Increased BL staff time by 32.5 hrs	22%	7%
Same teaching time	Increase BL staff time by 52.5 hrs	18%	3%
Same preparation and teaching time	Increase BL staff time by 85 hrs	15%	0%
BL staff time 20% higher than F2F	Increase BL staff time to 20% higher than F2F for preparation and teaching	8%	-9%
BL staff time 50% higher than F2F	Increase BL staff time to 50% higher than F2F for preparation and teaching	-4%	-23%

[a]Expressed as BL in relation to F2F. Positive values indicate BL is more cost-effective; negative values indicate F2F is more cost-effective.

Table 5 explores various transition cost scenarios and the resulting transition cost. Using the scenarios from Table 4, the analysis shows changes in 5-year accumulated PV as a result of modifying ongoing and transition costs. Negative values indicate that, at 5 years, the savings made from running the lower cost BL format have not yet overcome the transition cost. In fact, in the scenarios in the last 3 columns of the table, the cost of running the BL program is greater than or equal to that of the F2F program, and as such will never recover their transition costs.

Table 5. Multivariate sensitivity analysis adjusting transition costs and ongoing cost scenarios.

Transition scenario	Transition cost, AUD $	5-year accumulated PV difference in AUD $ by scenario (F2F - BL - transition cost)					
		Current model	Same preparation time	Same teaching time	Same preparation and teaching time	BL staff time 20% higher than F2F	BL staff time 50% higher than F2F
Current model	38,186	17,523	1,673	-22,653	-38,186	-87,655	-161,382
Added 210 hrs of IT support	50,931	4,779	-11,072	-35,398	-50,931	-100,400	-174,127
Break-even at 5 years	55,710	0	-15,851	-40,177	-55,710	-105,179	-178,906
Break-even at 10 years	101,499	-45,789	-61,640	-85,966	-101,499	-150,968	-224,695

XSL•FO
RenderX

Discussion

Principal Findings

This research aimed to determine the cost-effectiveness of an F2F and BL approach for EBM training within a medical program. The training had known outcomes evaluated through a sufficiently powered randomized controlled trial. The ICER results indicate that the BL approach provided an improved cost-effectiveness proposition from the perspective of the educational institution—costing less to train student clinicians to an equivalent level of competency. Cost-effectiveness results should be interpreted together with the break-even point of 3 years, as it is after this point that savings are realized by the institution. In addition to this, the findings by Ilic et al [12] indicated that the blended learning approach had the added benefit of increasing student attitudes toward EBM and self-reported use of EBM in clinical practice.

Given that blended learning can take a wide variety of formats, the findings of this study show that low-cost resources such as YouTube and student self-directed activities are cost-effective at improving EBM. This may be in part due to the nature of EBM teaching, as opposed to other aspects of medical education that may tend to utilize more costly mediums such as animations or virtual patients, which few medical schools can afford to create [31]. In fact, many costly features such as animations, high-quality video products, and excessive multimedia have little added value and may actually impede learning [20]. Additionally, given that the online learning utilized either currently owned or free access resources, there were no new costs attributed to software or licensing. Prior to using the BL approach in EBM, lecture material and online resources were already available through the online learning management system. In the RCT, the combination of using simple BL formats and staff having prior experience with learning technology may have helped to exclude transition costs such as IT support, consultation, and piloting. Despite this, the impact of support roles in transitioning learning format appears to be highly situation dependent, given that a study on developing an e-module found that approximately one third of development time was attributed to administrative or technical staff [19]. The impact of IT support costs on our cost model was tested in the sensitivity analysis, calculated as an additional 210 hours of staff time to transition. Despite significantly increasing the cost to transition, over 5 years, this cost was recovered by the savings in operating costs.

The differences in operating cost between the two approaches is dependent on staffing time, specifically the preparation and teaching time. The lower values in the BL approach can be attributed to greater emphasis on peer-to-peer learning and self-directed activities. Staff members act as facilitators rather than tutors, guiding students to resources rather than direct teaching. In contrast, previous studies have highlighted the significant increase in staffing time associated with online education methods [19,32]. Increases appear to be the result of using email, discussion forums, and online chat sessions, 75% of which occur outside regular working hours [19]. Thus, it is likely that not all BL formats would be cost-effective, and educators should carefully design their learning models around staff time, as this is the most significant cost driver. This is reflected in the sensitivity analysis showing that large increases in staffing hours above the traditional approach result in unfavorable ICER values and an inability to recover the transition cost. As would be expected, the change in teaching method had no impact on unit coordination, marking, or examination costs.

The BL format used in this study did not reduce the face-to-face classroom hours. Rather, there was an emphasis on how existing resources were used, how time was spent outside of class, and changing what activities were completed in class. That is, the cost reduction to achieve the same or better outcome observed in this study, is naturally a reflection of productivity gains around resources used based on enhanced teaching formats. As a result, the cost-effectiveness found in this approach may not be generalizable to institutions seeking to reduce student contact hours or promote distance education. Additionally, this study benefited from economies of scale as the EBM program educates approximately 500 students across multiple campuses. Small increases in education effectiveness are magnified when the effect is across a large number of students, resulting in favorable ICER and break-even results. This multiplication effect also applies to staff time, with a YouTube video produced by one person replacing the work of many people. If preparation time is reduced by just 20 minutes for a session, across 5 sites and 10 sessions the reduction in preparation adds up to over 15 hours. Thus, when designing a cost-effective program, it appears pertinent to consider not only what aspect of learning is being changed, but how many people the change will influence.

There may be concerns over the acceptability of a BL approach from a student or industry perspective. Social demands are driving information and communication technologies to be more commonplace and publically acceptable. It is conceivable that although the BL approach was accepted by students within this study, that this may have been a function of time, and the same approach may not have been acceptable in the preceding years. The concept of pre-class homework is not new, and changing the format does not necessarily solve the inherent issue of student adherence [20]. It remains unknown how different formats of BL will impact participation and outcomes [20]. Regardless, from the perspective of the student, pre-class activities require additional time and may be more or less efficient than traditional approaches [20]. To consider the student perspective on costs, a cost-benefit analysis using "willingness to pay" may be used. Such an analysis would help to consider student perceived value and inform changes in fee structure associated with changes in learning format.

Limitations

The generalizability of this study to other populations, as a result of the specific approach to BL used, have already been discussed. However, it is also necessary to discuss the limitations of the methodology used, which may influence the strength of the findings. The first limitation is the use of the ICER as a measure of cost-effectiveness. As mentioned, the standard use of the ICER does not include the investment cost of transitioning learning format. It is possible to include the investment cost

XSL•FO
RenderX

into the ICER calculation for the first iteration only, as this is the period when the cash flow occurs, when considering the investment as a prospective cost. However, the ICER would not be representative of other periods. By excluding the transition cost, the ICER presented in this study reflects the ongoing operational cost-effectiveness from the second iteration onwards and must be interpreted together with the break-even analysis. Thus, programs with high set-up costs and low running costs, which may be typical of many blended learning formats, are shown to be more cost-effective by the ICER. However, the blended format in the RCT had relatively low set-up costs, which have been roughly estimated by the course developers to be similar to the cost of transitioning from a blended approach to a traditional approach. For the same reasons, the BL 5-year major curriculum update was estimated to be the same as F2F. It is logical to infer that formats with high set-up costs will incur high updating costs; however, within an educational institution, updating is often limited by time availability. Another consideration with the ICER is that it assumes a linear increase in EBM expertise. That is, the assumption that an increase in Berlin score from 30% to 40% represents the same increase in EBM expertise as an increase from 70% to 80%. There is insufficient literature on the Berlin Questionnaire to draw conclusions on the proficiency distribution, and while this does not preclude the use of the ICER calculation, it should be considered when interpreting the results.

In the RCT itself, the quality of the teacher within both methods may also influence the costs and the educational outcomes. Within the study, this potential variable was controlled for by using the same teaching staff across both modes of education. However, in using the same staff, there is the risk that staff may inadvertently bias the results in favor of the teaching model that they are more skilled or enthusiastic about [33]. The scope of research does not include variations in the savings made by the institution. If the blended and face-to-face approaches were offered with different fee structures, the total value proposition from the educational institution will change. Less tangible benefits from the modes of delivery, such as feelings of connectedness often associated with face-to-face learning [34], flexibility [20], and other measures of the full educational experience have not been calculated, and remain as unknown for both of the educational approaches utilized [35-37]. And finally, there is the assumption that increasing the number of EBM trained practitioners, competent in incorporating best evidence, with patient values and clinical practice, actually makes a difference to patient outcomes.

No approach to cost analysis in medical education is perfect; however, being transparent about the approach used is an important step forward considering the current state of the literature. Here, we feel we have been as transparent as possible about the costs of the components that made up the various approaches to medical education described. The most common errors in cost analyses include omitting hidden costs and paying insufficient attention to the main costs. In this regard, we feel we have gone to considerable lengths to seek out and account for hidden costs and that we have paid most attention to the main costs—that is, the costs of faculty and resources.

Conclusions

Under the study settings, a blended learning approach to training practitioners to be competent in applying evidenced-based medicine was more cost-effective to operate than the traditional face-to-face model. Furthermore, the BL approach resulted in significantly greater increases in student attitudes toward EBM and self-reported use of EBM in clinical practice. When taking into account the cost of transitioning to the new format, the benefit of the cost-effectiveness is realized by the institution only after the third operational year. The primary drivers of cost-effectiveness were the low-cost online resources chosen, decreased staffing levels, and economies of scale. Implementing BL is not without its risks though and requires a significant investment cost in tailoring the teaching and learning resources to the Web-based environment during the transition to this approach. Using a BL approach will not necessarily be cost-effective, and consideration should be given to the blend utilized, staff expertise, and the educational setting. Health professions' education and educational research has developed into a respectable scientific discipline due to the shift toward scientific rigor and peer-review [38,39]. To maintain its relevance and accountability and to improve the adoption of new educational approaches and innovations, the next cultural shift in this field needs to be toward fiscal responsibility alongside learning outcomes, such as measuring outcomes of cost-effectiveness alongside measures of educational outcomes and the learning experience. The purpose of this shift is not to cut costs or to increase spending but simply to improve value.

Authors' Contributions

All authors have significantly contributed to the research within the manuscript, its analysis, and the writing and critical review of the final manuscript.

Conflicts of Interest

None declared.

References

1. Straus S, Glasziou P, Richardson W, Haynes B. Evidence-based medicine: how to practice and teach it. Edinburgh: Churchill Livingstone: Elsevier; 2011.
2. Finkel ML, Brown H, Gerber LM, Supino PG. Teaching evidence-based medicine to medical students. Med Teach 2003 Mar;25(2):202-204. [doi: 10.1080/0142159031000092634] [Medline: 12745533]

XSL•FO
RenderX

3. Meats E, Heneghan C, Crilly M, Glasziou P. Evidence-based medicine teaching in UK medical schools. Med Teach 2009 Apr;31(4):332-337. [doi: 10.1080/01421590802572791] [Medline: 19404893]
4. Ilic D, Hart W, Fiddes P, Misso M, Villanueva E. Adopting a blended learning approach to teaching evidence based medicine: a mixed methods study. BMC Med Educ 2013;13:169 [FREE Full text] [doi: 10.1186/1472-6920-13-169] [Medline: 24341502]
5. Nair B, Coughlan J, Hensley M. Student and patient perspectives on bedside teaching. Med Educ 1997;31(5):341-346. [doi: 10.1046/j.1365-2923.1997.00673.x]
6. Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ, Montori VM. Instructional design variations in internet-based learning for health professions education: a systematic review and meta-analysis. Acad Med 2010 May;85(5):909-922. [doi: 10.1097/ACM.0b013e3181d6c319] [Medline: 20520049]
7. Grasl MC, Pokieser P, Gleiss A, Brandstaetter J, Sigmund T, Erovic BM, et al. A new blended learning concept for medical students in otolaryngology. Arch Otolaryngol Head Neck Surg 2012 Apr;138(4):358-366. [doi: 10.1001/archoto.2012.145] [Medline: 22508619]
8. Lehmann R, Bosse H, Simon A, Nikendei C, Huwendiek S. An innovative blended learning approach using virtual patients as preparation for skills laboratory training: perceptions of students and tutors. BMC Medical Education 2013;13:23. [doi: 10.1186/1472-6920-13-23]
9. Woltering V, Herrler A, Spitzer K, Spreckelsen C. Blended learning positively affects students' satisfaction and the role of the tutor in the problem-based learning process: results of a mixed-method evaluation. Adv Health Sci Educ Theory Pract 2009 Dec;14(5):725-738. [doi: 10.1007/s10459-009-9154-6] [Medline: 19184497]
10. Ellaway R. E-learning: is the revolution over? Med Teach 2011;33(4):297-302. [doi: 10.3109/0142159X.2011.550968] [Medline: 21456987]
11. Maloney S, Tai J, Paynter S, Lo K, Ilic D. Self-Directed Online Learning Modules: Students' Behaviours and Experiences. Pharmacy 2013 Jul 29;1(1):8-15. [doi: 10.3390/pharmacy1010008]
12. Ilic D, Nordin R, Glasziou P, Tilson J, Villanueva E. A randomised controlled trial of a blended learning education intervention for teaching evidence-based medicine. BMC Medical Education 2015;15:39. [doi: 10.1186/s12909-015-0321-6]
13. Kremer M, Brannen C, Glennerster R. The challenge of education and learning in the developing world. Science 2013 Apr 19;340(6130):297-300. [doi: 10.1126/science.1235350] [Medline: 23599477]
14. Kremer M, Holla A. Improving Education in the Developing World: What Have We Learned from Randomized Evaluations? Annu Rev Econom 2009;1:513-542 [FREE Full text] [doi: 10.1146/annurev.economics.050708.143323] [Medline: 23946865]
15. Zurn P, Dal Poz Mario R, Stilwell B, Adams O. Imbalance in the health workforce. Hum Resour Health 2004 Dec 17;2(1):13 [FREE Full text] [doi: 10.1186/1478-4491-2-13] [Medline: 15377382]
16. Walsh K. In: Walsh K, editor. Cost Effectiveness in Medical Education. United Kingdom: Radcliffe Medical PR; 2010.
17. Walsh K, Reeves S, Maloney S. Exploring issues of cost and value in professional and interprofessional education. J Interprof Care 2014 Nov;28(6):493-494. [doi: 10.3109/13561820.2014.941212] [Medline: 25051088]
18. Maloney S, Haas R, Keating JL, Molloy E, Jolly B, Sims J, et al. Breakeven, cost benefit, cost effectiveness, and willingness to pay for web-based versus face-to-face education delivery for health professionals. J Med Internet Res 2012;14(2):e47 [FREE Full text] [doi: 10.2196/jmir.2040] [Medline: 22469659]
19. Delgaty L. A critical examination of the time and workload involved in the design and delivery of an e-module in postgraduate clinical education. Med Teach 2013 May;35(5):e1173-e1180. [doi: 10.3109/0142159X.2012.737963] [Medline: 23151211]
20. Cook DA. The value of online learning and MRI: finding a niche for expensive technologies. Med Teach 2014 Nov;36(11):965-972. [doi: 10.3109/0142159X.2014.917284] [Medline: 25072533]
21. Whitten P, Mair F, Haycox A, May C, Williams T, Hellmich S. Systematic review of cost effectiveness studies of telemedicine interventions. BMJ 2002 Jun 15;324(7351):1434-1437 [FREE Full text] [Medline: 12065269]
22. Upton D, Upton P. Development of an evidence-based practice questionnaire for nurses. J Adv Nurs 2006 Feb;53(4):454-458. [doi: 10.1111/j.1365-2648.2006.03739.x] [Medline: 16448488]
23. Huang G, Reynolds R, Candler C. Virtual patient simulation at US and Canadian medical schools. Acad Med 2007 May;82(5):446-451. [doi: 10.1097/ACM.0b013e31803e8a0a] [Medline: 17457063]
24. Stotzer RL, Fujikawa K, Sur J, Arnsberger P. Cost Analysis of a Distance Education MSW Program. Journal of Teaching in Social Work 2013 Nov;33(4-5):357-368. [doi: 10.1080/08841233.2013.826318]
25. Ilic D. Year 3-EBCP. 2013. URL: https://www.youtube.com/user/EBMOLT [accessed 2015-05-26] [WebCite Cache ID 6YoLQirHD]
26. Monash University. 2013. Evidence Based Practice URL: http://guides.lib.monash.edu/medicine [accessed 2015-05-26] [WebCite Cache ID 6YoM49Dhy]
27. Fritsche L, Greenhalgh T, Falck-Ytter Y, Neumayer H, Kunz R. Do short courses in evidence based medicine improve knowledge and skills? Validation of Berlin questionnaire and before and after study of courses in evidence based medicine. BMJ 2002 Dec 7;325(7376):1338-1341 [FREE Full text] [Medline: 12468485]
28. Philips C, Thompson G. What is cost-effectiveness?. 2003. URL: http://www.medicine.ox.ac.uk/bandolier/painres/download/whatis/Cost-effect.pdf [accessed 2015-05-26] [WebCite Cache ID 6YoMM3KWv]

XSL•FO
RenderX

29. World Health Organization. Making choices in health: WHO guide to cost-effectiveness analysis. Geneva: World Health Organization; 2003. URL: http://www.who.int/choice/publications/p_2003_generalised_cea.pdf [accessed 2015-07-16] [WebCite Cache ID 6a4TZoO12]
30. Monash University. 2014. Monash University Enterprise Agreements URL: http://adm.monash.edu.au/enterprise-agreements/ [accessed 2015-05-26] [WebCite Cache ID 6YoMExAWL]
31. Huang G, Reynolds R, Candler C. Virtual patient simulation at US and Canadian medical schools. Acad Med 2007 May;82(5):446-451. [doi: 10.1097/ACM.0b013e31803e8a0a] [Medline: 17457063]
32. Spector JM. Time demands in online instruction. Distance Education 2005 Jan;26(1):5-27. [doi: 10.1080/01587910500081251]
33. Devereaux P, Bhandari M, Clarke M, Montori V, Cook D, Yusuf S, et al. Need for expertise based randomised controlled trials. BMJ 2005 Jan 8;330(7482):88 [FREE Full text] [doi: 10.1136/bmj.330.7482.88] [Medline: 15637373]
34. Besser H, Bonn M. Impact of distance independent education. J Am Soc Inf Sci 1996 Nov;47(11):880-883. [doi: 10.1002/(SICI)1097-4571(199611)47:11<880::AID-ASI14>3.0.CO;2-Z]
35. Wells R. Computer-mediated Communication for Distance Education: An International Review of Design, Teaching, and Institutional Issues. Pennsylvania: American Center for the Study of Distance Education; 1992.
36. Phelps RH, Wells RA, Ashworth RL, Hahn HA. Effectiveness and costs of distance education using computer - mediated communication. American Journal of Distance Education 1991 Jan;5(3):7-19. [doi: 10.1080/08923649109526759]
37. Rumble B. Online costs: interactivity at a price. In: Mayson R, Kaye A, editors. Communication, Computers and Distance Education. Oxford: Pergamon Press; 1989.
38. Feuer MJ, Towne L, Shavelson RJ. Scientific Culture and Educational Research. Educational Researcher 2002 Nov 01;31(8):4-14. [doi: 10.3102/0013189X031008004]
39. Levin J. What to do about educational research's credibility gaps? Issues in Education 1999;5(2):177-229. [doi: 10.1016/S1080-9724(00)00025-2]

Abbreviations

BL: blended learning
EMB: evidence-based medicine
F2F: face-to-face
ICER: incremental cost-effectiveness ratio
MD: mean difference
PV: present value
QASE: quality-adjusted students educated
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 12.02.15; peer-reviewed by S Sadasivan, D Cook; comments to author 23.03.15; accepted 08.07.15; published 21.07.15.

Please cite as:
Maloney S, Nicklen P, Rivers G, Foo J, Ooi YY, Reeves S, Walsh K, Ilic D
A Cost-Effectiveness Analysis of Blended Versus Face-to-Face Delivery of Evidence-Based Medicine to Medical Students
J Med Internet Res 2015;17(7):e182
URL: http://www.jmir.org/2015/7/e182/
doi:10.2196/jmir.4346
PMID:26197801

©Stephen Maloney, Peter Nicklen, George Rivers, Jonathan Foo, Ying Ying Ooi, Scott Reeves, Kieran Walsh, Dragan Ilic. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 21.07.2015. 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

Letter to the Editor

Photoaging Mobile Apps: A Novel Opportunity for Smoking Cessation?

Titus J Brinker[1*]; Werner Seeger[1*]

Universities of Giessen and Marburg Lung Center (UGMLC), Member of the German Center for Lung Research (DZL), Justus-Liebig-University Giessen, Giessen, Germany

[*]all authors contributed equally

Corresponding Author:
Titus J Brinker
Universities of Giessen and Marburg Lung Center (UGMLC)
Member of the German Center for Lung Research (DZL)
Justus-Liebig-University Giessen
Klinkstr. 32
Giessen, 35392
Germany
Phone: 49 15175084347
Fax: 49 15175084347
Email: titus.brinker@gmail.com

(J Med Internet Res 2015;17(7):e186) doi:10.2196/jmir.4792

KEYWORDS

tobacco prevention, tobacco cessation, photoaging, mobile phone apps, apps, app, smartphones

Most smokers start smoking during their early adolescence with the idea that smoking is glamorous; the problems related to lung cancer, vascular disease, and chronic pulmonary disease are too far in the future to fathom. In contrast, most adolescents view their image in a mirror as an important component of their personal life. A recent randomized controlled trial by Burford et al published in the *Journal of Medical Internet Research* demonstrated an increased quit rate of 21% in 18-30-year-old young adults by the help of photoaging desktop programs, in which an image is altered to predict future appearance [1]. Furthermore, the photoaging software has been shown to increase the motivation of 14-18-year-old females to quit [2]. However, the investigated programs only reach a small audience and are not freely available.

We took advantage of the widespread availability of mobile phones and adolescents´ interest in appearance to develop a free mobile phone app which requires the user to take a self portrait (ie, a selfie), which is then displayed by the photoaging software as four images: consequences of (non-)smoking one pack a day for a year (Figure 1) or 15 years (Figure 2). Afterwards, the app explains the visual results and offers many sharing options with family and friends. By this means, the social network of the user may also be informed about the various beauty reducing effects of smoking, potential health consequences, and learn about the app.

The underlying aging algorithms take into account the user's current age and are based on publications showing an increased risk for acne and pale skin due to declined capillary perfusion (after one pack-year), as well as connective tissue changes and wrinkles in the longer term (after 15 pack-years) [3,4].

The app has been installed on over 50,000 Android and 27,000 iOS mobile phones within seven months after its release in Germany (10/27/2014 to 4/26/2015). As mobile phone use in Germany declines with age, the largest fraction of the app's users are assumed to be 30 years or younger.

Based on the publication from Burford et al, it is reasonable to speculate that the app could motivate smokers to quit. Taking into account that the smoking prevalence in the general German population is approximately 25% (approximately 19,250 of the 77,000 app users were smokers), about 4000 users (21%) would have quit after using the app. Further research is needed to investigate the effectiveness of app-based photoaging interventions to increase quit rates and to prevent smoking initiation.

XSL•FO
RenderX

Figure 1. Photoaged image of a 17 year old woman showing the consequences of smoking one pack a day for one year (vs. non-smoking).

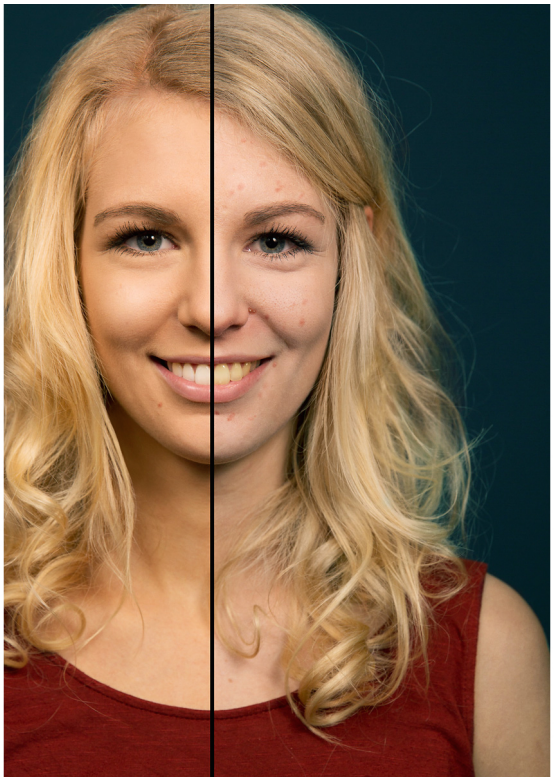
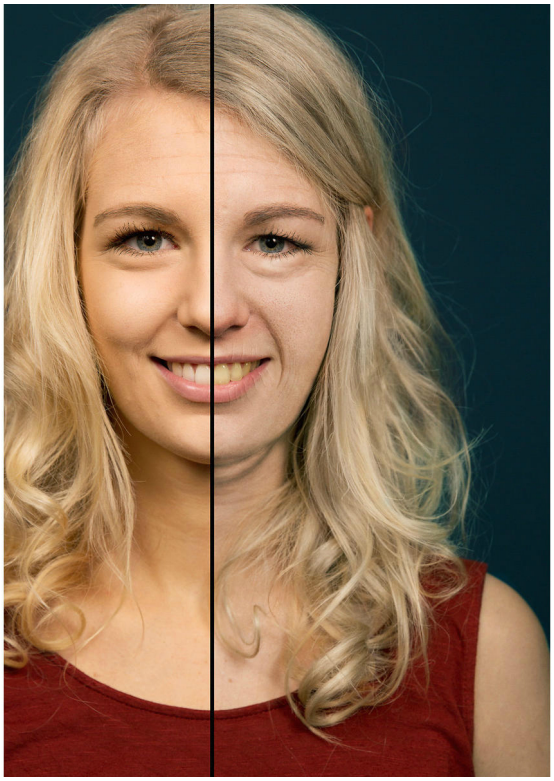


Figure 2. Photoaged image of a 17 year old woman showing the consequences of smoking one pack a day for 15 years (vs non-smoking).



Acknowledgments

Titus J Brinker holds a Kaltenbach-scholarship and a grant on tobacco research; both from the German Heart Foundation.

References

XSL•FO
RenderX

1. Burford O, Jiwa M, Carter O, Parsons R, Hendrie D. Internet-based photoaging within Australian pharmacies to promote smoking cessation: randomized controlled trial. J Med Internet Res 2013;15(3):e64 [FREE Full text] [doi: 10.2196/jmir.2337] [Medline: 23531984]
3. Weiss C, Hanebuth D, Coda P, Dratva J, Heintz M, Stutz EZ. Aging images as a motivational trigger for smoking cessation in young women. Int J Environ Res Public Health 2010 Sep;7(9):3499-3512 [FREE Full text] [doi: 10.3390/ijerph7093499] [Medline: 20948939]
3. Grady D, Ernster V. Does cigarette smoking make you ugly and old? Am J Epidemiol 1992 Apr 15;135(8):839-842. [Medline: 1585896]
4. Schäfer T, Nienhaus A, Vieluf D, Berger J, Ring J. Epidemiology of acne in the general population: the risk of smoking. Br J Dermatol 2001 Jul;145(1):100-104. [Medline: 11453915]

Edited by G Eysenbach; submitted 02.06.15; peer-reviewed by O Burford; comments to author 19.06.15; accepted 07.07.15; published 27.07.15.

Please cite as:
Brinker TJ, Seeger W
Photoaging Mobile Apps: A Novel Opportunity for Smoking Cessation?
J Med Internet Res 2015;17(7):e186
URL: http://www.jmir.org/2015/7/e186/
doi:10.2196/jmir.4792
PMID:26215210

©Titus J Brinker, Werner Seeger. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 27.07.2015. 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