Journal of Medical Internet Research

Journal Impact Factor (JIF) (2022): 7.4
Volume 15 (2013), Issue 1 ISSN 1438-8871 Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Original Papers

Original Paper

Effectiveness of a Web-Based Cognitive-Behavioral Tool to Improve Mental Well-Being in the General Population: Randomized Controlled Trial

John Powell[1], MBBChir, MSc, PhD, FFPH; Thomas Hamborg[2], PhD; Nigel Stallard[2], PhD; Amanda Burls[1], MBBS, MSc, FFPH; Jaime McSorley[3]; Kylie Bennett[4], BSc, BA(Hons); Kathleen M Griffiths[4], BSc(Hons), PhD; Helen Christensen[5], PhD

[1]Department of Primary Care Health Sciences, University of Oxford, Oxford, United Kingdom
[2]Division of Health Sciences, Warwick Medical School, University of Warwick, Coventry, United Kingdom
[3]NHS Choices, Department of Health, London, United Kingdom
[4]Centre for Mental Health Research, The Australian National University, Canberra, Australia
[5]Black Dog Institute, University of New South Wales, Randwick, Australia

Corresponding Author:
John Powell, MBBChir, MSc, PhD, FFPH
Department of Primary Care Health Sciences
University of Oxford
23-38 Hythe Bridge Street
Oxford, OX1 2ET
United Kingdom
Phone: 44 01865617855
Fax: 44 01865617855
Email: john.powell@phc.ox.ac.uk

Abstract

Background: Interventions to promote mental well-being can bring benefits to the individual and to society. The Internet can facilitate the large-scale and low-cost delivery of individually targeted health promoting interventions.

Objective: To evaluate the effectiveness of a self-directed Internet-delivered cognitive-behavioral skills training tool in improving mental well-being in a population sample.

Methods: This was a randomized trial with a waiting-list control. Using advertisements on a national health portal and through its mailing list, we recruited 3070 participants aged 18 or over, resident in England, and willing to give their email address and access a fully automated Web-based intervention. The intervention (MoodGYM) consisted of 5 interactive modules that teach cognitive-behavioral principles. Participants in the intervention arm received weekly email reminders to access the intervention. The control group received access to the intervention after the trial was completed and received no specific intervention or email reminders. Outcomes were assessed by using self-completion questionnaires. The primary outcome was mental well-being measured with the Warwick-Edinburgh Mental Well-being Scale (WEMWBS). Secondary outcomes were Center for Epidemiologic Studies Depression scale (CES-D) depression scores, Generalized Anxiety Disorder 7-item scale (GAD-7) anxiety scores, EuroQol Group 5-Dimension Self-Report Questionnaire (EQ-5D) quality of life scores, physical activity, and health service use. All outcomes were measured at baseline, and at 6- and 12-week follow-ups.

Results: A total of 1529 (49.80%) participants completed final follow-up at 12 weeks. Retention was 73.11% (1123/1536) in the control arm and 26.47% (406/1534) in the intervention arm. No relationship between baseline measures and withdrawal could be established. The analysis of WEMWBS mental well-being scores using a linear mixed model for repeated measures showed no difference between intervention and control group at baseline (difference –0.124 points, 95% CI –0.814 to 0.566), and significant improvements for the intervention group at 6 weeks (2.542 points, 95% CI 1.693-3.390) and at 12 weeks (2.876 points, 95% CI 1.933-3.819). The model showed a highly significant (*P*<.001) intervention by time interaction effect. There were also significant improvements in self-rated scores of depression and anxiety. Given the high level of attrition, a sensitivity analysis with imputed missing values was undertaken that also showed a significant positive effect of the intervention.

Conclusions: Participants allocated to the intervention arm had an average increase of approximately 3 points on the WEMWBS scale compared to no increase for participants in the control group. Three points on this scale is approximately one-third of a standard deviation. In a low-cost automated intervention designed to shift the population distribution of mental well-being, a small difference per individual could yield a major benefit in population terms. In common with other Web-based interventions, there were high rates of attrition. Further work is needed to improve acceptability, to evaluate against placebo effect, and to disaggregate the effect on mental well-being from the effect on depression and anxiety.

Trial Registration: International Standard Randomised Controlled Trial Number Register ISRCTN 48134476; http://www.controlled-trials.com/ISRCTN48134476 (Archived by WebCite® at http://www.webcitation.org/6DFgW2p3Q)

(J Med Internet Res 2013;15(1):e2) doi:10.2196/jmir.2240

KEYWORDS

Mental health; Public health; Randomized controlled trial; Internet

Introduction

Interventions to promote positive mental health and well-being can bring benefits both to the individual by improving mood and psychological functioning, and also to society in terms of economic prosperity and social cohesion [1]. There is now worldwide interest in the promotion of mental well-being with measures of well-being being adopted as key economic indicators alongside gross domestic product (GDP) [2]. Yet there are few studies of individually targeted interventions with a primary aim of promoting mental well-being. In theory, an approach using the principles of cognitive behavioral therapy (CBT) to encourage more healthy patterns of thinking and behavior may offer an individual-level intervention to promote positive mental health. There is evidence for the effectiveness of CBT approaches in preventing depression (primarily among adolescents and young adults) [3-5], in improving resilience (often group interventions delivered in workplace settings) [6], and in the promotion of workplace well-being yielding benefits such as improvements in productivity, sickness absence, and stress [7]. There is also an emergent literature on the promotion of well-being using positive psychology interventions that encompass a range of psychological approaches including cognitive-behavioral aspects [8].

At the same time, the Internet is playing an increasingly important role in health care. It can provide a platform for the large-scale delivery of information and interventions for modifying lifestyle risk factors that result in more informed and empowered citizens who are better able to manage their own health. The area of e-mental health has been of particular interest to researchers and practitioners [9] because online tools, such as Internet-delivered computerized cognitive behavioral therapies (CCBT), have been shown to be effective for a range of mental health conditions [10], both when combined with therapist contact and when fully automated [11]. The Internet is also being used for positive psychology approaches [12]. The fully automated Web-based MoodGYM intervention was originally developed as a tool to prevent depression in young people and has been demonstrated to be effective in this context [13]. It has also been shown to be acceptable, safe, effective, and cost-effective in alleviating symptoms of mild to moderate depression and anxiety in community samples [14-17]. Although self-directed Internet interventions are known to have low rates of adherence [18], this is less of a problem in well-being promotion for the general population than for the treatment of mental illness because it does not raise ethical questions of inadequacy of treatment for a diagnosed health problem. Moreover, as a mental health promotion tool, the intervention can be delivered at very low marginal cost by using minimal personnel resources so that it can be made freely available to all who wish to use it, in contrast to a therapist contact approach that is neither feasible nor affordable for all.

In this study, we undertook a randomized controlled trial to test the effectiveness of a Web-based individually targeted self-help CBT package (MoodGYM) for promoting mental well-being in the general population.

Methods

Study Design and Participants

We undertook a randomized trial with two parallel group arms: intervention and a waiting-list control. Recruitment took place over 2 weeks in September 2010. Participants were self-recruited users of the UK National Health Service (NHS) NHS Choices website who were invited to take part in an online trial to promote mental well-being. Self-completion pop-up user surveys conducted previously showed that in 2010 most users of this NHS Choices website were women (76%), and 68% of users were in the 25 to 64 years age range. To be eligible for our study, participants were required to confirm that they were aged 18 or over, lived in England (as covered by our ethics and governance approval), and had Internet access and an email address.

Procedures

Study recruitment advertisements were placed on the NHS Choices website (specifically the Live Well and mental health pages), in the NHS Choices newsletter sent to all subscribers (approximately 80,000), in emails sent to NHS Choices Customer Insights research group, and on the NHS Choices Facebook and Twitter pages, as well as on the Carers Direct Facebook page. These advertisements offered participants the opportunity to take part in a mental fitness trial, with the aim of promoting mental well-being. The study was not advertised as a treatment for people who were ill; the emphasis was on mental health promotion. Those interested in participating were invited to complete an online form to confirm eligibility, and to read information about the study. After a period of 48 hours to allow them time to reflect on their decision to take part in the research, eligible participants were invited by email to provide informed consent via an online form, to create a username and password, and to complete baseline questionnaires. Trial administration was automated and participants remained quasi-anonymous, identified only by email address. Multiple registrations by single email address were forbidden.

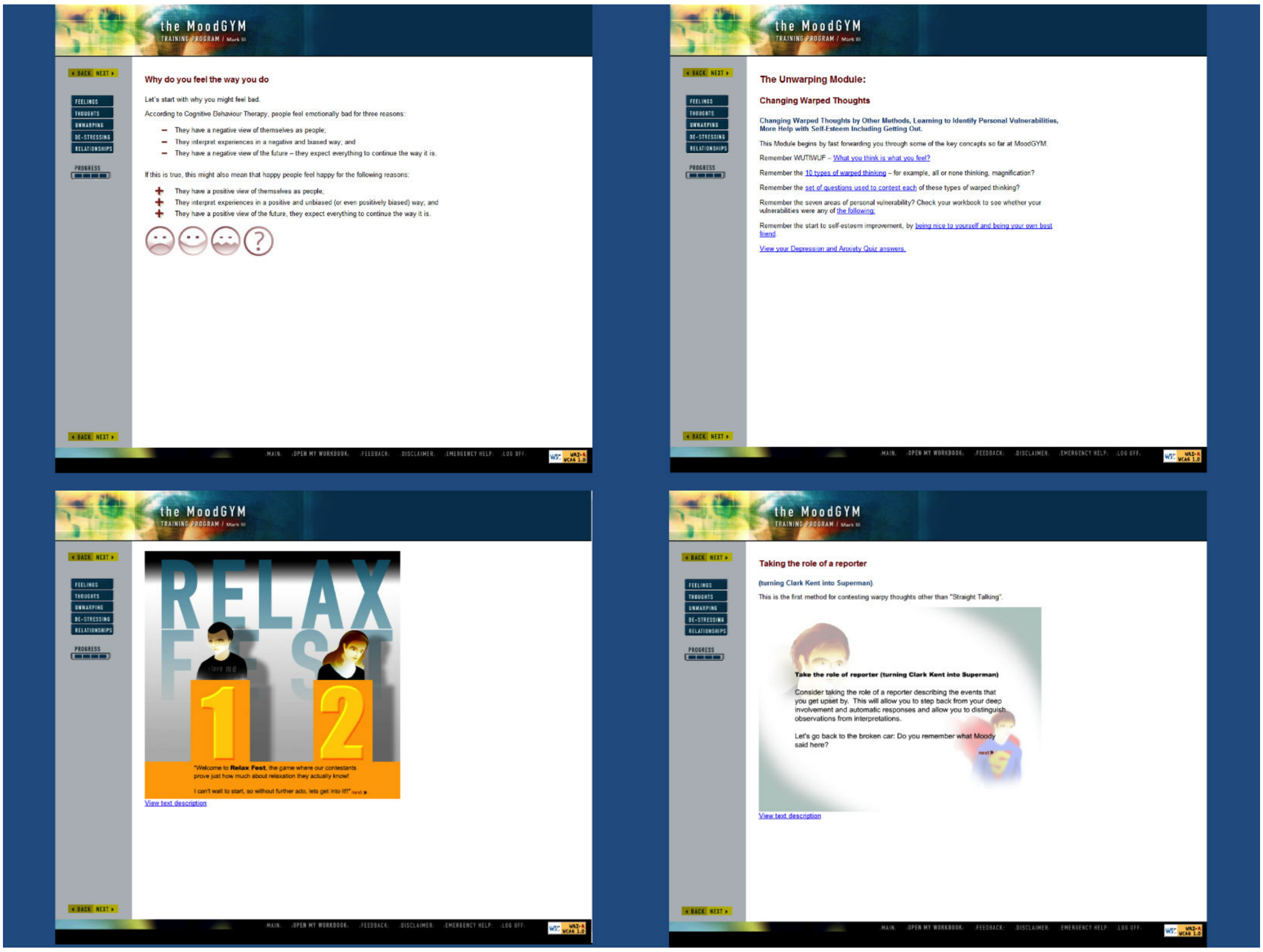
After completion of baseline measures, participants were sent an automated email directing them to log in to a trial portal with their new username and password. At this point participants were automatically randomized to either the intervention or control group. Once randomized, participants were immediately provided with access to the intervention (intervention group) or they were given general information about accessing the NHS Choices Healthy Living pages and informed that they would receive the intervention after a period of 3 months (waiting-list control group). Randomization was in a 1:1 ratio using predefined automated computerized block randomization with a block size of 2. The automated computerized system was set up by technical staff not involved in the day-to-day management of the study. Allocation was concealed from the researchers. As we chose to use a waiting-list control, participants were not blind to whether or not they were in the intervention group. To prevent contamination in the control arm, we did not use the name “MoodGYM” in the study documentation. Participants were free to withdraw at any time without giving a reason. There were no content changes, periods of downtime, or bug fixes required during the trial.

We received approvals from the NHS ethics committee (Black Country REC 10/H1202/21), the Australian National University (ANU) Human Research Ethics Committee (protocol number 2010/244) and NHS research governance. The study was registered on the International Standard Randomised Controlled Trial Number Register (ISRCTN 48134476).

Intervention

MoodGYM is a free Internet-based self-help program that teaches cognitive-behavioral skills. It consists of 5 interactive modules that use diagrams and online exercises. It demonstrates the relationship between thoughts and emotions, examines issues related to stress and to relationships, and teaches relaxation and meditation techniques. It also includes sections on managing relationships and problem solving. Screenshots of the MoodGYM intervention are shown in Figure 1. Participants are encouraged to work their way through each of the 5 modules, 1 module per week, but are able to work at their own pace, ad libitum. The program includes an online workbook with 29 online exercises to help promote mental health. We made slight modifications to some phrases used in the MoodGYM tool to replace Australian colloquialisms with their English equivalent. We added logos to indicate affiliation to the NHS and University of Warwick (lead academic institution). Participants in the intervention arm received weekly email reminders to log in to the trial portal where they could access the intervention.

Figure 1. Screenshots of MoodGYM intervention.



Why do you feel the way you do

Let's start with why you might feel bad.

According to Cognitive Behaviour Therapy, people feel emotionally bad for three reasons:

- They have a negative view of themselves as people;
- They interpret experiences in a negative and biased way; and
- They have a negative view of the future – they expect everything to continue the way it is.

If this is true, this might also mean that happy people feel happy for the following reasons:

+ They have a positive view of themselves as people;
+ They interpret experiences in a positive and unbiased (or even positively biased) way; and
+ They have a positive view of the future, they expect everything to continue the way it is.

The Unwarping Module:

Changing Warped Thoughts

Changing Warped Thoughts by Other Methods, Learning to Identify Personal Vulnerabilities, More Help with Self-Esteem Including Getting Out.

This Module begins by fast forwarding you through some of the key concepts so far at MoodGYM.

Remember WUTIWUF – What you think is what you feel?

Remember the 10 types of warped thinking – for example, all or none thinking, magnification?

Remember the set of questions used to contest each of these types of warped thinking?

Remember the seven areas of personal vulnerability? Check your workbook to see whether your vulnerabilities were any of the following:

Remember the start to self-esteem improvement, by being nice to yourself and being your own best friend.

View your Depression and Anxiety Quiz answers.

"Welcome to **Relax Fest**, the game where our contestants prove just how much about relaxation they actually know!

I can't wait to start, so without further ado, lets get into it!!"

View text description

Taking the role of a reporter

(turning Clark Kent into Superman).

This is the first method for contesting warpy thoughts other than "Straight Talking".

Take the role of reporter (turning Clark Kent into Superman)

Consider taking the role of a reporter describing the events that you get upset by. This will allow you to step back from your deep involvement and automatic responses and allow you to distinguish observations from interpretations.

Let's go back to the broken car: Do you remember what Moody said here?

View text description

Control Group

The comparator was a waiting-list control group. During the trial, the control participants did not receive any specific intervention or email reminders. In common with the participants in the intervention arm, control participants were able to access general information pages on mental well-being on the NHS Choices website. At the completion of the trial (3 months after its commencement), participants in the control group were provided with access to the intervention.

Outcome Measures

The primary outcome measure was mental well-being as measured using the self-completion Warwick-Edinburgh Mental Well-being Scale (WEMWBS) [19]. This 14-item instrument has been validated for the UK population and adopted by the Scottish Health Survey and the Health Survey for England. It asks respondents to read statements about feelings and thoughts and to choose the response (a 5-point scale ranging from none of the time to all of the time) that best describes their experience over the previous two weeks. Example items are "I've been feeling optimistic about the future" and "I've been thinking clearly." It has been shown to have good content validity and shows high correlations with other scales of mental health and well-being. It has a near-normal population distribution, with no ceiling effects.

Secondary outcomes were self-completed Center for Epidemiologic Studies Depression scale (CES-D) depression scores, Generalized Anxiety Disorder 7-item (GAD-7) anxiety scores, EuroQol Group 5-Dimension Self-Report Questionnaire (EQ-5D) quality of life scores, physical activity (self-reported frequency of exercise), and use of health services (self-reported general practitioner consultations or hospital visits). All outcomes were measured at the start of the trial (baseline before the intervention), immediately following the intervention (6 weeks after baseline), and 6 weeks after the intervention was finished (12 weeks after baseline).

Statistical Analysis

The study was powered to detect a difference of 2 points in the change over time (to the 12-week endpoint) of the WEMWBS score. Based on an estimated population mean score of 49.8 (from the Scottish Health Survey 2008) [20], and a standard deviation for mean change in WEMWBS scores over time of 9.84 [21], we required approximately 510 participants in each group with full data (for 2-sided type I error rate $\alpha = .05$, power of 90%). Allowing for a high level of attrition (estimate 50%) as is common in fully automated Internet interventions, we aimed to recruit 2040 participants in total.

For the analysis of the primary endpoint and secondary endpoints where possible, (generalized) linear mixed models for repeated measures were fitted. These models appropriately account for the correlation between measurements from the same subject at different time points (baseline, 6-week follow-up, and 12-week follow-up). Models for each endpoint consisted of 3 effects: measurement occasion (time), intervention (MoodGYM or waiting-list control), and the interaction effect of time and intervention. Of primary interest was the intervention by time interaction effect. This effect informs whether the intervention type had a differential effect on the change over time in the two groups, thus answering the primary research hypothesis. Adjusted least squares means estimates and standard errors are presented for each endpoint and each model. An unstructured covariance matrix was used for modeling of correlations between repeated observations as this covariance matrix yielded the best fit among investigated structures for all endpoints.

Secondary endpoints which did not satisfy distributional assumptions for the repeated measures analysis were compared using paired *t* tests. These *t* tests were utilized to compare changes of outcome values between time points for MoodGYM and waiting-list control rather than absolute outcome values. Change scores fulfilled distributional assumptions of the *t* test where applied. Simple descriptive statistics (mean, median, standard deviation, range) were used to compare baseline characteristics of the two groups. The statistical analysis was conducted using the statistical software package SAS release 9.2 (SAS Institute, Inc, Cary, NC, USA). Ordinary linear mixed models were fitted using the MIXED procedure and the GLIMMIX was employed for fitting generalized linear mixed models. A 2-sided type I error rate of 5% was used throughout.

Analyses were conducted on an intention-to-treat basis, including all participants in the groups to which they were randomized.

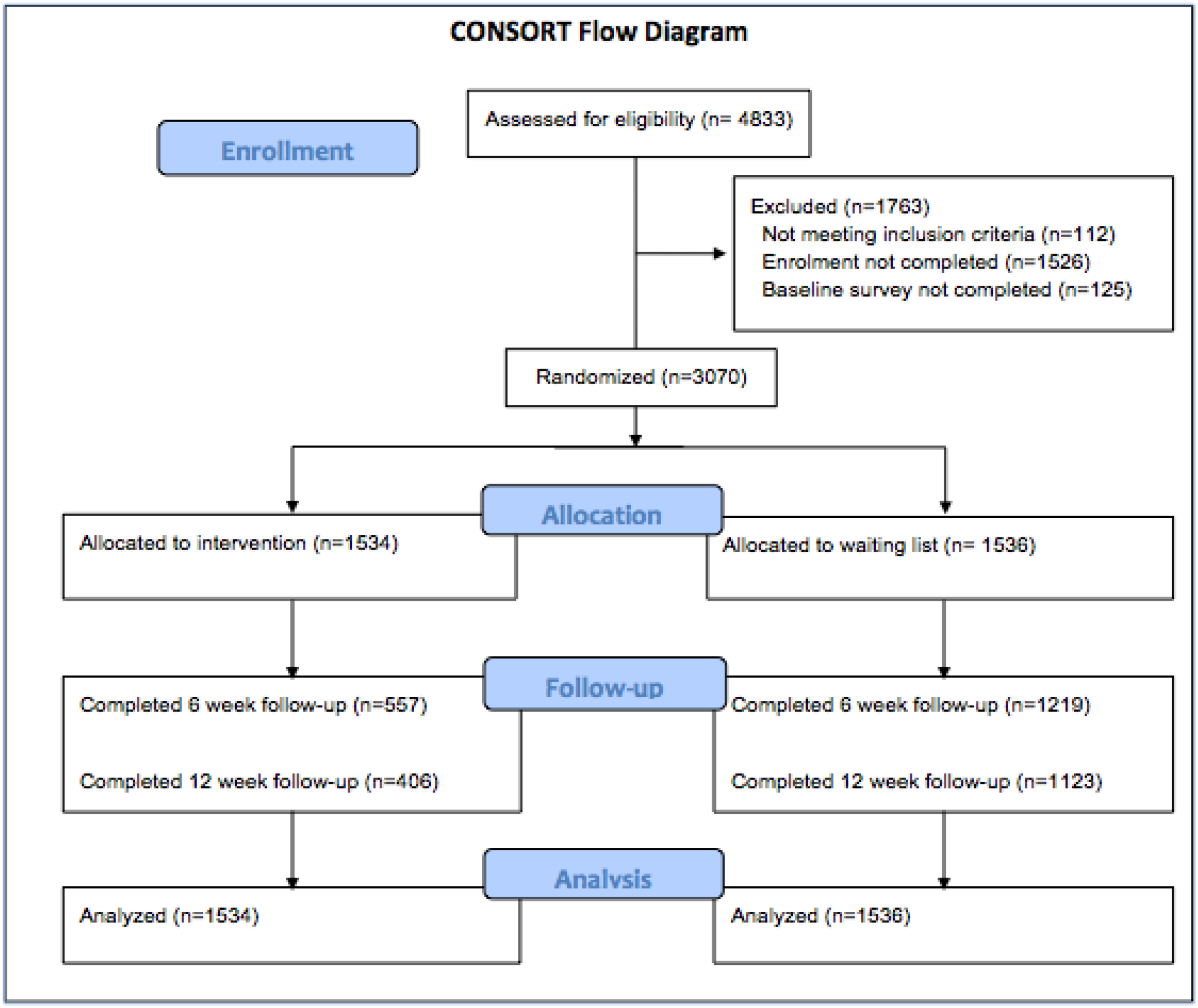
Results

Participation Rates

The trial flow diagram (Figure 2) shows participant recruitment and retention at baseline, 6-week follow-up, and 12-week follow-up. Over a 2-week period in September 2010, 8589 people accessed the URL; 4833 people completed the eligibility screening for the study and were sent invitation emails. Of these, 3070 returned the completed consent forms and baseline measures and were randomized into the study.

Attrition rate was high in this study. Total losses to follow-up were 50%, with participants in the intervention arm more likely to withdraw from the study. Attrition was 73.5% in the intervention arm and 26.9% in the control arm (risk ratio 2.76, 95% CI 2.53-3.02). No relationship between baseline characteristics and likelihood of withdrawal could be established. The WEMWBS score at baseline and posttest was slightly higher for participants retained in the trial until the end for both arms than for those participants who withdrew, but this difference was small and not statistically significant. A small number of participants (61 MoodGYM, 77 control) had missing observations at the 6-week follow-up, but provided responses at the later follow-up.

Figure 2. CONSORT flow diagram.



Comparison of Baseline Characteristics

Table 1 summarizes the demographics and baseline characteristics of participants in the trial who were well balanced between the treatment groups. Most of participants were female (77.88%), in line with the general profile of the users of the NHS Choices portal that we used for recruitment, and 92.15% reported white ethnicity. The mean age was 41 years. Most participants were using the Internet daily and rated themselves as having either good or excellent Internet ability. More than half of the study participants had previously received treatment for a mental health problem.

Table 1. Demographic and baseline characteristics of trial participants (N=3070).

Variable		MoodGYM (n=1534)	Control (n=1536)
Gender, n (%)			
	Female	1179 (76.86)	1212 (78.91)
Age (years), mean (SD)		40.88 (12.96)	41.39 (13.05)
Ethnicity, n (%)			
	White	1408 (91.79)	1421 (92.51)
	Mixed	23 (1.50)	26 (1.69)
	Asian	33 (2.15)	30 (1.95)
	Black	26 (1.69)	16 (1.04)
	Other	29 (1.89)	31 (2.02)
	Invalid/missing	15 (0.98)	12 (0.78)
Marital status, n (%)			
	Married/cohabiting	723 (47.13)	711 (46.29)
	Divorced/separated	266 (17.34)	292 (19.01)
	Never married	545 (35.53)	533 (34.70)
Employment status, n (%)			
	Working	953 (62.13)	916 (59.64)
	Student	56 (3.65)	61 (3.97)
	Retired	93 (6.06)	101 (6.58)
	Looking after home/family	202 (13.17)	234 (15.23)
	Unemployed	174 (11.34)	181 (11.78)
	Other	56 (3.65)	43 (2.80)
Smoking, n (%)			
	Daily	218 (14.21)	214 (13.93)
	Occasionally	103 (6.71)	108 (7.03)
Units of alcohol in past week, mean (SD)		2.93 (3.88)	3.14 (4.16)
Drug use in past week, n (%)			
	Yes	53 (3.46)	40 (2.60)
Internet use frequency, n (%)			
	At least once a day	1361 (88.72)	1348 (87.76)
	Several times week	165 (10.76)	171 (11.13)
	Less than once a week	8 (0.52)	17 (1.11)
Internet ability, n (%) [a]			
	Excellent	847 (55.22)	822 (53.52)
	Good	560 (36.51)	578 (37.63)
	Fair	123 (8.02)	127 (8.27)
	Poor	2 (0.13)	4 (0.26)
	Bad	2 (0.13)	1 (0.07)
General health score, mean (SD)		69.12 (21.10)	69.04 (20.40)
Previous treatment of a mental health problem, n (%)			
	Yes	877 (57.17)	843 (54.88)

Variable	MoodGYM (n=1534)	Control (n=1536)
Previous CBT experience, n (%)		
Yes	326 (21.25)	321 (20.90)
Previous Internet-based CBT, n (%)		
Yes	116 (7.56)	114 (7.42)
Number of days in past week with > 30 minutes physical activity, n (%)		
0	418 (27.25)	408 (26.56)
1	254 (16.56)	261 (16.99)
2	287 (18.71)	281 (18.29)
3	222 (14.47)	221 (14.39)
4	134 (8.74)	107 (6.97)
5	99 (6.45)	124 (8.07)
6	36 (2.35)	48 (3.13)
7	84 (5.48)	86 (5.60)

[a] Control group responses n=1532.

Primary Endpoint Analysis

The primary research hypothesis of the trial was that MoodGYM improves well-being measured by WEMWBS at 6-week and 12-week follow-ups. Table 2 displays the adjusted WEMWBS score means on each measurement occasion. The difference at baseline, 6 weeks, and 12 weeks was –0.124 (95% CI –0.814 to 0.566), 2.542 (95% CI 1.693-3.390), and 2.876 (95% CI 1.933-3.819) points, respectively.

Table 2. Estimates of marginal means over balanced populations and standard errors for the Warwick-Edinburgh Mental Well-being Scale (WEMWBS).

Time point	WEMWBS scores	
	MoodGYM	Control
Baseline		
Marginal mean	42.20	42.32
Standard error	0.251	0.246
6-week follow-up		
Marginal mean	44.46	41.92
Standard error	0.343	0.265
12-week follow-up		
Marginal mean	45.17	42.30
Standard error	0.387	0.285

The results from a mixed model repeated measures analysis including time, intervention, and the interaction between them are given in the upper part of Table 3. The interaction effect (intervention × time point) is highly significant (*P*<.001), indicating that the intervention, MoodGYM, has a differential treatment effect compared with the control arm. A partition of the interaction effect to provide comparison of the two groups at each time point [22], indicated that there was no difference at baseline (*P*=.72) but that differences at 6 and 12 weeks were both highly significant (*P*<.001 in each case). Covariates were added to the model individually to determine those that had an influence on the model fit. All covariates that improved the model fit (in terms of the Akaike Information Criterion [AIC] using a likelihood ratio test) were included in the full model shown in Table 3. The overall model fit was significantly better than for the model excluding covariates (*P*<.001). The intervention by time interaction remained highly significant (*P*<.001). Although previous treatment for a mental health problem explained a significant amount of variation in the model, previous treatment did not have a significant impact on the change of WEMWBS scores over the study duration. There was no significant covariate by intervention interaction for any of the investigated covariates.

Table 3. Type III test of fixed effects for primary endpoint (WEMWBS).

Effect	*F* test (*df*)	*P*
Basic model (AIC = 44,482) [a]		
Intervention	23.74 (1,3068)	< .001
Time point	25.68 (2,3301)	< .001
Intervention × time point	33.87 (2,3301)	< .001
Full model (AIC = 43,959) [a]		
Intervention	31.08 (1,3057)	< .001
Time point	25.44 (2,3299)	< .001
Intervention × time point	33.51 (2,3299)	< .001
Mental health service	197.04 (1,3057)	< .001
Mental health service × time point	0.46 (2,3299)	.63
Physical activity	16.00 (7,3057)	< .001
Previous CBT use	19.92 (1,3057)	< .001
Smoking	24.11 (2,3057)	< .001

[a] AIC: Akaike Information Criterion.

To explore whether bias was introduced through systematic participant dropout, we undertook a completer analysis using the observed mean scores for those participants who completed all 3 investigations. Mean scores for those who adhered to the follow-up schedule were slightly higher on each occasion in both the MoodGYM and control arms than for the overall study population. However, mean WEMWBS scores at baseline for completers were only 0.42 and 0.13 points above the full study population means for MoodGYM and control groups, respectively, indicating that there was no systematic dropout of participants with lower baseline scores.

Given the high level of attrition, a sensitivity analysis with imputed missing values was conducted to evaluate the robustness of the effect of MoodGYM on mental well-being [23]. Missing values were imputed using the last observation carried forward (LOCF) procedure. This procedure is known to possess poor properties, underestimating variability and producing biased treatment effect estimates [24]. The LOCF procedure was only used here as a sensitivity analysis to investigate robustness given the high level of attrition. The estimated mean WEMWBS scores for the imputed dataset at 12-week follow-up were 43.34 and 42.24 for MoodGYM and control, respectively. Even under this highly conservative assumption, the interaction of intervention and time remained highly significant (P<.001).

Analysis of Secondary Endpoints

Depression and Anxiety

The results presented in Tables 4 and 5 show that MoodGYM and the waiting-list control had a significantly different effect on the CES-D score and GAD-7 score over time with participants in the intervention (MoodGYM) arm reporting a reduction in levels of depression and anxiety. The time by intervention interaction was highly significant (P<.001) for both endpoints. The differences at baseline, 6 weeks, and 12 weeks were –0.041 (95% CI –0.993 to 0.911), –2.793 (95% CI –3.947 to –1.640), and –3.365 (95% CI –4.621 to –2.110) points respectively for CES-D, and 0.212 (95% CI –0.074 to 0.758), –1.124 (95% CI –1.607 to –0.642), and –1.495 (95% CI –2.030 to –0.960) points for GAD-7.

Table 4. Estimates of marginal means over balanced populations and standard errors for Center for Epidemiologic Studies Depression scale (CES-D) endpoint.

Time point		CES-D scores	
		MoodGYM	Control
Baseline			
	Marginal mean	23.23	23.27
	Standard error	0.338	0.348
6-week follow-up			
	Marginal mean	20.38	23.17
	Standard error	0.469	0.356
12-week follow-up			
	Marginal mean	19.30	22.67
	Standard error	0.515	0.381

Table 5. Estimates of marginal means over balanced populations and standard errors for Generalized Anxiety Disorder 7-item scale (GAD-7) endpoint.

Time point		GAD-7 scores	
		MoodGYM	Control
Baseline			
	Marginal mean	8.80	8.46
	Standard error	0.151	0.149
6-week follow-up			
	Marginal mean	7.17	8.29
	Standard error	0.193	0.153
12-week follow-up			
	Marginal mean	6.60	8.10
	Standard error	0.221	0.161

Quality of Life, Physical Activity, and Health Service Use

The EQ-5D quality of life data were bimodal rendering an analysis using a linear mixed model impossible. Instead, *t* tests comparing the individual changes of EQ-5D scores between the two treatment arms were conducted. Change data were sufficiently normally distributed for the changes between baseline and the second follow-up measurement. There were no significant differences between arms for the change to 6 weeks (*P*=.78) or to 12 weeks (*P*=.42).

Physical activity was measured by using an ordered categorical variable ("In the past week, on how many days have you done a total of 30 minutes or more of physical activity, which was enough to raise your breathing rate"). Therefore, we used a mixed-effects proportional odds model to analyze this endpoint. This models the probability of being in category "*x* days of activity" or fewer days of activity for each of the 8 categories. Direct maximum likelihood estimation was used instead of the restricted maximum likelihood estimation employed in presented linear mixed models. The interaction term was significant indicating that the patterns of change of physical activity over time were significantly different between the two groups. Although significant, the effect (*F* test) was smaller than for the primary outcome and the secondary endpoints CES-D and GAD-7. The impact of time appears to be greater than the impact of treatment on this endpoint. Examination of the data suggests that the difference is explained by participants in the control group being more likely to report reduced activity at 12 weeks.

The number of general practitioner visits and hospital outpatient visits during the previous month were reported at baseline, 6 weeks, and 12 weeks and compared between the 2 groups. Repeated measures generalized linear mixed model for count data assuming a Poisson distribution were fitted for both secondary endpoints. The results showed that there was no significantly different effect between the groups on the mean number of GP visits over time (*P*=.30). There was also no differential effect between the MoodGYM and waiting-list control groups on the mean number of hospital attendances (*P*=.32).

Prespecified Subgroup Analyses

In order to investigate the consistency of the treatment effect, subgroup analyses based on age, gender, psychiatric history, previous use of CBT, level of anxiety, and level of depression were prespecified in the protocol. For each subgroup, a mixed model consisting of a time, group, and time × group effect was fitted. The results of these analyses are shown in Table 6. The treatment effect was very consistent across subgroups. Changes in the WEMWBS primary outcome remained significant for all subgroups except those aged under 26 years (change of 2.13 points, *P*=.22); however, the numbers in this subgroup were small as indicated by the wide confidence intervals. The 95% CI for participants over the age of 25 is entirely contained in the CI for those aged under 26 years. Thus, there is no evidence for a differential treatment effect between these two subgroups. Of note, the nondepressed and nonanxious subgroups of those with a CES-D score less than 16, or a GAD-7 score less than 10, both showed significant improvement in their WEMWBS well-being scores (*P*<.001).

Table 6. Subgroup analysis for primary endpoint, the Warwick-Edinburgh Mental Well-being Scale (WEMWBS) score.

Subgroup characteristic	Baseline			12-week follow-up				*F* (*df*)[a]	*P*
	MoodGYM mean (n)	Control mean (n)	Mean diff	MoodGYM mean (n)	Control mean (n)	Mean diff	95% CI		
Age < 26 years	41.93 (182)	41.51 (172)	0.42	45.63 (30)	43.50 (114)	2.13	–0.94, 5.20	1.50 (2,319)	.22
Age > 25 years	42.23 (1352)	42.42 (1364)	–0.19	45.13 (376)	42.17 (1009)	2.96	1.97, 3.95	33.40 (2,2978)	<.001
Female	42.29 (1179)	42.27 (1212)	0.02	44.95 (327)	42.15 (889)	2.80	1.76, 3.83	23.08 (2,2614)	<.001
Male	41.88 (355)	42.50 (324)	–0.62	45.97 (79)	42.80 (234)	3.17	0.97, 5.37	11.75 (2,683)	<.001
Psychiatric history	39.47 (877)	39.85 (843)	–0.38	42.48 (219)	39.59 (621)	2.89	1.62, 4.17	20.97 (2,1830)	<.001
No psychiatric history	45.83 (657)	45.32 (693)	0.51	48.74 (187)	45.59 (502)	3.15	1.83, 4.47	13.51 (2,1467)	<.001
Previous CBT	38.70 (326)	39.40 (321)	–0.70	41.00 (93)	38.74 (251)	2.26	0.37, 4.16	7.40 (2,750)	.007
No previous CBT	43.14 (1208)	43.09 (1215)	0.05	46.38 (313)	43.26 (872)	3.12	2.06, 4.19	28.18 (2,2547)	<.001
GAD-7 < 10	46.70 (915)	46.77 (934)	–0.07	48.77 (242)	46.00 (683)	2.78	1.67, 3.88	17.72 (2,2034)	<.001
GAD-7 > 9	35.53 (619)	35.42 (602)	0.11	39.90 (164)	36.60 (440)	3.30	1.83, 4.77	16.94 (2,1263)	<.001
CES-D<16	51.46 (506)	51.51 (514)	–0.05	51.99 (139)	49.72 (379)	2.28	0.85, 3.70	9.70 (2,1114)	<.001
CES-D>15	37.64 (1028)	37.70 (1022)	–0.06	41.76 (267)	38.57 (744)	3.18	2.09, 4.28	24.71 (2,2183)	<.001
CES-D>26	33.94 (601)	33.96 (596)	–0.02.	38.52 (150)	35.74 (437)	2.77	1.31, 4.24	15.82 (2,1265)	<.001

[a] group × time.

Use of Intervention

Figure 3 shows the number of completed modules of the MoodGYM intervention by number of participants in the intervention arm.

A post hoc exploratory dose-response analysis to investigate the relationship between number of modules of the MoodGYM intervention completed and change in well-being, revealed that the change from baseline WEMWBS score to score at 12 weeks was significant for participants in the intervention group ("condition=moodgym") who completed 2 or more modules (Figure 4).

Figure 3. Number of completed modules by participants in intervention group.

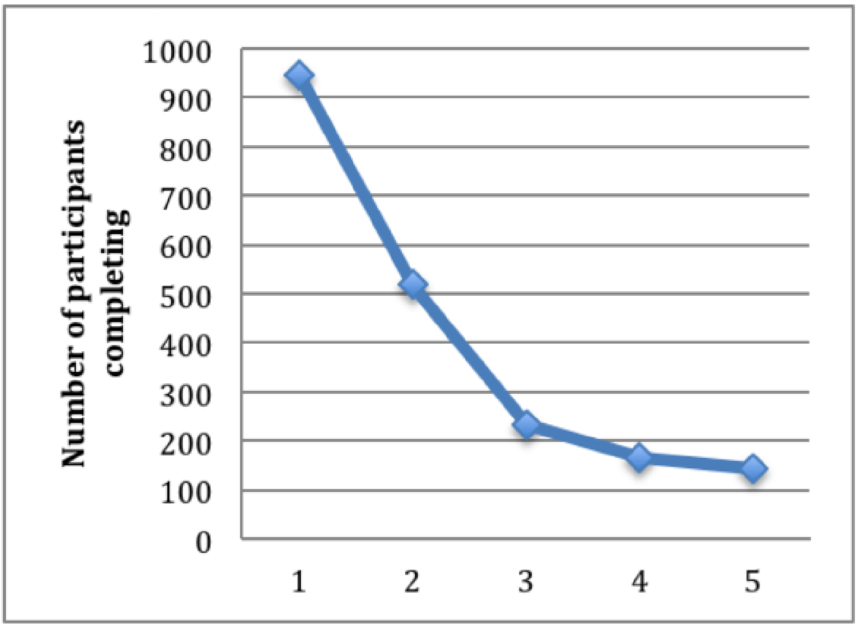
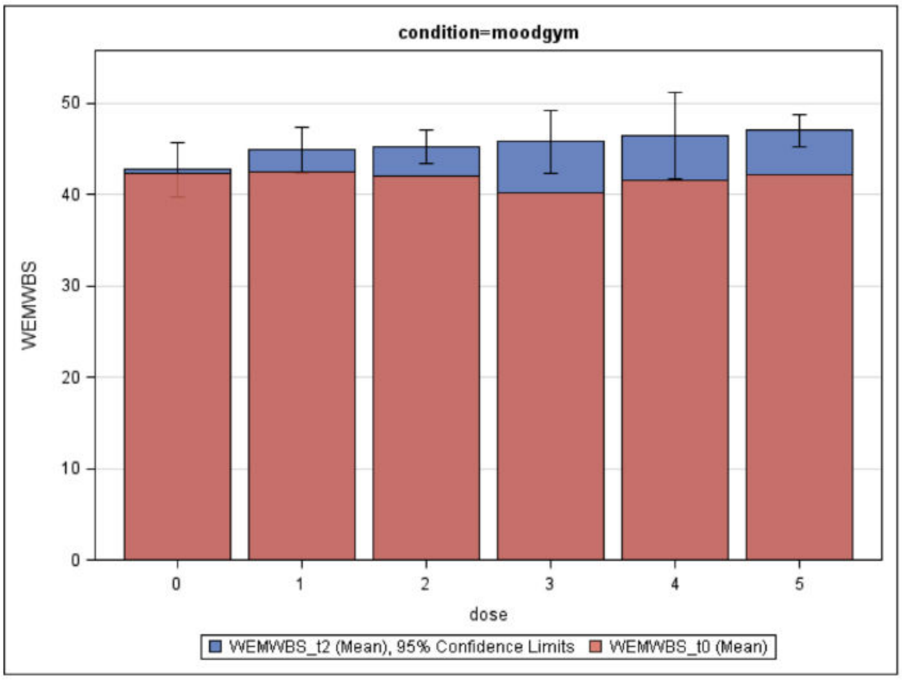


Figure 4. Mean Warwick-Edinburgh Mental Well-being Scale (WEMWBS) scores by number of completed weeks. Red is mean score at baseline and blue is mean score at 12-week follow-up.



Adverse Events

We received emails from 2 participants who indicated that they were suffering some level of mental distress that could possibly be related to the intervention. Neither was deemed to be a serious adverse event by the Trial Steering Committee, and both were reported to the ethics committee. In the first instance, a trial participant said that they no longer wished to continue with the trial having found one section of the intervention (on warpy thoughts) difficult to complete. In the second instance, a trial participant reported finding the intervention distressing to complete and asked to be withdrawn. Both participants were withdrawn immediately and given advice on seeking help from their primary care provider or from mental health services.

Discussion

Main Findings

We successfully delivered a fully automated health promoting intervention to a large sample of the general population using the Internet. On average, those allocated to receive the intervention improved their mental well-being scores by almost 3 points on the WEMWBS scale over a 12-week period, whereas the scores for those in the waiting-list control group (who received no intervention) remained nearly unchanged. This effect was highly statistically significant (*P*<.001). The observed change of 2.876 points on the WEMWBS scale represents an effect size (Cohen’s *d*) of approximately 0.34. In a public health intervention designed to shift the whole distribution of mental well-being upwards in a population, such a difference can be important because a small difference per individual can bring a major benefit in population terms (as seen, for example, in public health interventions to reduce blood pressure). Analyses of secondary outcomes showed significant improvements (*P*<.001) in self-report measures of depression (CES-D) and anxiety (GAD-7). There were no significant differences in measures of quality of life (EQ-5D) or self-reported health service use. There was also a significant difference (*P*=.002) in self-reported physical activity at 12-week follow-up, explained by participants in the control group being more likely to report reduced activity. Our data on participant usage confirms high attrition rates and shows that a relatively low proportion of participants completed all 5 modules, and a post hoc dose-response analysis found statistically significant improvements in mental well-being (from baseline scores) in those completing 2 or more modules.

Limitations

Although we sought volunteers from the general population, the people who volunteered to take part in the research had relatively low initial mental well-being scores, which is not surprising given that we requested volunteers to take part in research to improve their mental well-being and the recruitment routes included advertisements placed on the mental health webpages of NHS Choices. The mean WEMWBS score for our participants was 42. The general population average (obtained from the Scottish Health Survey) is 49.8 (SD 8.3) [20]. Our study population also had relatively high mean scores on measures of depression (CES-D scale) and anxiety (GAD-7) scale, and a high level of previous treatment of a mental health problem, confirming that although our aim was to recruit volunteers from across the general population, an intervention for promoting mental well-being had particular salience for those with some level of mental health problems. This means we cannot be certain that a similar increase in well-being would be observed in a population with no prior mental health problems, although, importantly, our subgroup analyses showed that the treatment effect remained highly significant between the two arms of the trial within the subgroup of nondepressed participants and within the subgroup of participants with no previous treatment of mental health problems.

The trial was waiting-list controlled, so we cannot rule out the possibility of a placebo effect. We did not follow up participants beyond 3 months and further work on long-term effectiveness would be desirable. There was a low level of male participation in the trial, although the ratio of male to female participants was in line with the profile of users of the portal from which we recruited. There was a high level of dropout from the trial, particularly in the intervention arm. We tried to minimize dropout by incorporating a 48-hour period between passing eligibility screening and being accepted into the trial and by not randomizing until after all baseline measures had been completed. In this way, we hoped to recruit participants with some commitment to returning to the website and participating in the study. Most people who dropped out did not inform us, but simply stopped returning to the site or responding to emails. It is likely that more participants were retained in the control arm as they had an incentive to stay in (they were on the waiting list to receive the intervention), and the tasks they were required to complete during the trial (surveys at 6 weeks and 12 weeks) were less demanding than the intervention group (intervention and surveys). Importantly, no systematic differences between those who dropped out and those who completed in either the intervention or the control groups could be identified, and there was no systematic dropout of participants with lower baseline scores. A sensitivity analysis that assumed that those who dropped out would have had no change in their well-being scores, showed that under this assumption the intervention would still have had a significant positive effect on mental well-being. Self-directed Internet interventions are known to have low rates of adherence [25], but this is potentially less of a problem in well-being promotion for the general population than for the treatment of mental illness because it does not raise ethical questions of inadequacy of treatment of a diagnosed health problem. Moreover, as a mental health promotion tool, the intervention can be delivered at very low marginal cost using minimal personnel resources so that it can be made freely available to all who wish to use it, in contrast to a therapist contact approach which would be neither feasible nor affordable for all.

Comparison With Other Studies

This was the first trial to evaluate the promotion of mental well-being using an Internet-based CBT approach. Previous trials of Internet-based CBT approaches have shown effectiveness in treating mild to moderate depression and in the prevention of depression [11]. A recent systematic review found 5 randomized controlled trials that used positive psychology interventions (PPI), some of which used cognitive-behavioral principles, delivered over the Internet and measured well-being as an outcome [26]. Three of the studies in the review targeted adults with depression [27-29]. Of the 2 studies that included general population samples, one used a strengths intervention (identifying and using your strengths) delivered to an Australian population recruited through online advertisements. This trial showed a significant improvement on 1 of 4 a priori well-being outcome measures (the Personal Well-being Index), but not on the other 3 [30]. The other trial tested an online resilience-training package for sales managers, also in Australia that found no improvement on the Authentic Happiness Index [31]. Both of these trials had high levels of attrition (83% and 41.5%, respectively) as found in our study. Both studies had far fewer participants than in the present study (160 and 53 participants, respectively). The study by Mitchell and colleagues [26] used an information-only placebo control group given online information about problem solving, whereas the study by Abbott and colleagues [31] used a waiting-list control group.

Conclusions

This study demonstrated that a low-cost, easily accessible, highly scalable, and self-directed intervention delivered in a fully automated fashion can be effective at improving mental well-being among regular Internet users recruited from the general population accessing a national health portal in England. Given the potential societal benefits of an increase in population well-being and the cost advantages of Internet-delivery with no practitioner contact, this could have major implications if accessed more widely. We have also demonstrated in this study that a national health portal provides a feasible and acceptable platform for the successful and rapid recruitment of participants into research. The trial procedures including consent and all baseline and follow-up measures were fully automated with implications for the future conduct and cost of trials with designs that could harness this.

Further work is needed to evaluate the effect of MoodGYM on mental well-being against a control website, to follow up participants more completely and for longer periods of time, and to target those who are not currently depressed to disaggregate the effect on mental well-being from the effect on depression and anxiety. This last aim could perhaps be achieved by recruiting participants from a nonhealth website. There is also a general need to further explore the relationship between intervention adherence and outcomes [32]. Intervention development could follow other investigators in this field and explore how to increase adherence, perhaps by examining user motivations to persist [33], and trialing alternate modes of delivery [34], which may include mobile health applications. Finally, there is also a need for rigorous evaluation of CBT-based approaches in comparison with other approaches that may improve well-being, such as positive psychology and mindfulness interventions.

Acknowledgments

We thank Professor Mike Clarke, the independent chair of our Trial Steering Group; Bob Gann (Head of Strategy and Engagement) and Helen Rowntree (Strategy, Engagement and Digital Media Director) at NHS Choices; Anthony Bennett, e-hub IT manager at ANU, who provided technical expertise to the trial design and administration; and Ada Tam, e-hub Project Officer at ANU, who undertook user management duties. We also thank the participants for taking part in the trial.

JP, HC, KG, and KB conceived the study. JP wrote the original protocol with contributions from HC, KG, KB, AB, and NS. JP was chief investigator and oversaw the study. KB coordinated the trial administration at ANU. JM project managed the trial. TH undertook the statistical analysis. NS oversaw the conduct of the analysis. JP, AB, HC, and KG interpreted the findings. All authors contributed to the final manuscript. All authors had full access to all data. JP is guarantor for the paper.

The research was funded by NHS Choices. NHS Choices provided the electronic platform for participant recruitment. MoodGYM was provided by The Australian National University who delivered the intervention and administered the trial. KG is supported by NHMRC Fellowship no. 525413.

Conflicts of Interest

JP and JM worked for NHS Choices (JP as part-time Clinical Director, JM as project manager) who funded this study and provided the platform for participant recruitment. HC and KG are the authors of the MoodGYM intervention evaluated in this study. HC, KG, and KB work for The Australian National University who provide free access to MoodGYM on their website. TH, NS, and AB have no financial or nonfinancial interests to declare in relation to this study.

Multimedia Appendix 1

Completed CONSORT-EHEALTH checklist V.16 [35].

[PDF File (Adobe PDF File), 486KB - jmir_v15i1e2_app1.pdf]

Multimedia Appendix 2

Trial protocol as requested in CONSORT-EHEALTH [35].

[PDF File (Adobe PDF File), 352KB - jmir_v15i1e2_app2.pdf]

Multimedia Appendix 3

Trial information for participants as requested in CONSORT-EHEALTH [35].

[PDF File (Adobe PDF File), 117KB - jmir_v15i1e2_app3.pdf]

References

1. Slade M. Mental illness and well-being: the central importance of positive psychology and recovery approaches. BMC Health Serv Res 2010;10:26 [FREE Full text] [doi: 10.1186/1472-6963-10-26] [Medline: 20102609]
2. Stiglitz J, Sen A, Fitoussi JP. Commission on the Measurement of Economic Performance and Social Progress. Paris: Centre de Recherche en Economié de Sciences; 2009. Report by the Commission on the Measurement of Economic Performance and Social Progress URL: http://www.stiglitz-sen-fitoussi.fr/documents/rapport_anglais.pdf [accessed 2012-12-22] [WebCite Cache ID 6D6VHR963]
3. Venning A, Kettler L, Eliott J, Wilson A. The effectiveness of Cognitive-Behavioural Therapy with hopeful elements to prevent the development of depression in young people: a systematic review. Int J Evid Based Healthc 2009 Mar;7(1):15-33. [doi: 10.1111/j.1744-1609.2009.00122.x] [Medline: 21631843]
4. Barrera AZ, Torres LD, Muñoz RF. Prevention of depression: the state of the science at the beginning of the 21st Century. Int Rev Psychiatry 2007 Dec;19(6):655-670. [doi: 10.1080/09540260701797894] [Medline: 18092243]
5. Richardson T, Stallard P, Velleman S. Computerised cognitive behavioural therapy for the prevention and treatment of depression and anxiety in children and adolescents: a systematic review. Clin Child Fam Psychol Rev 2010 Sep;13(3):275-290. [doi: 10.1007/s10567-010-0069-9] [Medline: 20532980]
6. Proudfoot JG, Corr PJ, Guest DE, Dunn G. Cognitive-behavioural training to change attributional style improves employee well-being, job satisfaction, productivity, and turnover. Personality and Individual Differences 2009;46(2):147-153. [doi: 10.1016/j.paid.2008.09.018]

7. Graveling RA, Crawford JO, Cowie H, Amati C, Vohra S. National Institute for Health and Clinical Excellence. Edinburgh: Institute of Occupational Medicine; 2008. A review of workplace interventions that promote mental well-being in the workplace. Draft report URL: http://www.nice.org.uk/nicemedia/pdf/MentalWellbeingWorkFinalReport.pdf [accessed 2012-12-22] [WebCite Cache ID 6D6VSAU9s]
8. Sin NL, Lyubomirsky S. Enhancing well-being and alleviating depressive symptoms with positive psychology interventions: a practice-friendly meta-analysis. J Clin Psychol 2009 May;65(5):467-487. [doi: 10.1002/jclp.20593] [Medline: 19301241]
9. Riper H, Andersson G, Christensen H, Cuijpers P, Lange A, Eysenbach G. Theme issue on e-mental health: a growing field in internet research. J Med Internet Res 2010;12(5):e74 [FREE Full text] [doi: 10.2196/jmir.1713] [Medline: 21169177]
10. Griffiths KM, Christensen H. Review of randomized controlled trials of Internet interventions for mental disorders and related conditions. Clinical Psychologist 2006;10(1):16-29. [doi: 10.1080/13284200500378696]
11. Griffiths KM, Farrer L, Christensen H. The efficacy of internet interventions for depression and anxiety disorders: a review of randomised controlled trials. Med J Aust 2010 Jun 7;192(11 Suppl):S4-11. [Medline: 20528707]
12. 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]
13. Calear AL, Christensen H, Mackinnon A, Griffiths KM, O'Kearney R. The YouthMood Project: a cluster randomized controlled trial of an online cognitive behavioral program with adolescents. J Consult Clin Psychol 2009 Dec;77(6):1021-1032. [doi: 10.1037/a0017391] [Medline: 19968379]
14. Christensen H, Griffiths KM, Jorm AF. Delivering interventions for depression by using the internet: randomised controlled trial. BMJ 2004 Jan 31;328(7434):265 [FREE Full text] [doi: 10.1136/bmj.37945.566632.EE] [Medline: 14742346]
15. Christensen H, Griffiths KM, Mackinnon AJ, Brittliffe K. Online randomized controlled trial of brief and full cognitive behaviour therapy for depression. Psychol Med 2006 Dec;36(12):1737-1746. [doi: 10.1017/S0033291706008695] [Medline: 16938144]
16. Mackinnon A, Griffiths KM, Christensen H. Comparative randomised trial of online cognitive-behavioural therapy and an information website for depression: 12-month outcomes. Br J Psychiatry 2008 Feb;192(2):130-134 [FREE Full text] [doi: 10.1192/bjp.bp.106.032078] [Medline: 18245031]
17. 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]
18. Neil AL, Batterham P, Christensen H, Bennett K, Griffiths KM. Predictors of adherence by adolescents to a cognitive behavior therapy website in school and community-based settings. J Med Internet Res 2009;11(1):e6 [FREE Full text] [doi: 10.2196/jmir.1050] [Medline: 19275982]
19. Tennant R, Hiller L, Fishwick R, Platt S, Joseph S, Weich S, et al. The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. Health Qual Life Outcomes 2007;5:63 [FREE Full text] [doi: 10.1186/1477-7525-5-63] [Medline: 18042300]
20. Corbett J, Dobbie F, Doig M, D'Souza J, Given L, Gray L, et al. The Scottish Government. Edinburgh: The Scottish Government; 2010. Scottish Health Survey 2009 - Volume 1: Main report URL: http://www.scotland.gov.uk/Resource/Doc/325403/0104975.pdf [accessed 2012-12-22] [WebCite Cache ID 6D6Vf6gSW]
21. Lindsay G, Band S, Cullen MA, Cullen S, Strand S, Hasluck C, et al. Department for Education. Coventry, UK: University of Warwick; 2008. Parenting Early Intervention Pathfinder Evaluation. Research Report DCSF-RW054 URL: https://www.education.gov.uk/publications/eOrderingDownload/DCSF-RW054.pdf [accessed 2012-12-22] [WebCite Cache ID 6D6Vs3RrB]
22. Winer BJ. Statistical Principles in Experimental Design, 2nd ed. New York: McGraw-Hill; 1991.
23. White IR, Horton NJ, Carpenter J, Pocock SJ. Strategy for intention to treat analysis in randomised trials with missing outcome data. BMJ 2011;342:d40 [FREE Full text] [Medline: 21300711]
24. Altman DG. Missing outcomes in randomized trials: addressing the dilemma. Open Med 2009;3(2):e51-e53 [FREE Full text] [Medline: 19946393]
25. 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]
26. Mitchell J, Vella-Brodrick D, Klein B. Positive psychology and the internet: A mental health opportunity. EJournal of Applied Psychology 2010;6(2):30-41 [FREE Full text]
27. Seligman ME, Steen TA, Park N, Peterson C. Positive psychology progress: empirical validation of interventions. Am Psychol 2005;60(5):410-421. [doi: 10.1037/0003-066X.60.5.410] [Medline: 16045394]
28. Shapira LB, Mongrain M. The benefits of self-compassion and optimism exercises for individuals vulnerable to depression. The Journal of Positive Psychology 2010;5(5):377. [doi: 10.1080/17439760.2010.516763]
29. Parks-Sheiner AC. Penn Libraries ScholaryCommons: Repository: Dissertations. Philadelphia, PA: University of Pennsylvania; 2009. Positive psychotherapy: building a model of empirically supported self-help URL: http://repository.upenn.edu/dissertations/AAI3363580/ [accessed 2012-12-22] [WebCite Cache ID 6D6WBed6f]
30. Mitchell J, Stanimirovic R, Klein B, Vella-Brodrick D. A randomized controlled trial of a self-guided internet intervention promoting well-being. Computers in Human Behavior 2009;25:749-760. [doi: 10.1016/j.chb.2009.02.003]

31. Abbott JA, Kaldo V, Klein B, Austin D, Hamilton C, Piterman L, et al. A cluster randomised trial of an internet-based intervention program for tinnitus distress in an industrial setting. Cogn Behav Ther 2009 Sep;38(3):162-173. [doi: 10.1080/16506070902763174] [Medline: 19675959]
32. 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]
33. Donkin L, Glozier N. Motivators and motivations to persist with online psychological interventions: a qualitative study of treatment completers. J Med Internet Res 2012;14(3):e91 [FREE Full text] [doi: 10.2196/jmir.2100] [Medline: 22743581]
34. Mohr DC, Duffecy J, Jin L, Ludman EJ, Lewis A, Begale M, et al. Multimodal e-mental health treatment for depression: a feasibility trial. J Med Internet Res 2010;12(5):e48 [FREE Full text] [doi: 10.2196/jmir.1370] [Medline: 21169164]
35. 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

AIC: Akaike Information Criterion
CBT: cognitive behavioral therapy
CCBT: computerized cognitive behavioral therapies
CES-D: Center for Epidemiologic Studies Depression scale
EQ-5D: EuroQol Group 5-Dimension Self-Report Questionnaire
GAD-7: Generalized Anxiety Disorder 7-item scale
LOCF: last observation carried forward
NHS: National Health Service
PPI: positive psychology interventions
WEMWBS: Warwick-Edinburgh Mental Well-being Scale

Edited by G Eysenbach; submitted 26.06.12; peer-reviewed by Y Khazaal; comments to author 19.09.12; revised version received 02.10.12; accepted 04.10.12; published 09.01.13.

Please cite as:
Powell J, Hamborg T, Stallard N, Burls A, McSorley J, Bennett K, Griffiths KM, Christensen H
Effectiveness of a Web-Based Cognitive-Behavioral Tool to Improve Mental Well-Being in the General Population: Randomized Controlled Trial
J Med Internet Res 2013;15(1):e2
URL: http://www.jmir.org/2013/1/e2/
doi:10.2196/jmir.2240
PMID:23302475

Original Paper

Website Usage and Weight Loss in a Free Commercial Online Weight Loss Program: Retrospective Cohort Study

Kevin O Hwang[1,2], MD, MPH; Jing Ning[3], PhD; Amber W Trickey[1,2], PhD; Christopher N Sciamanna[4], MD, MPH

[1]Department of Internal Medicine, The University of Texas Medical School at Houston, Houston, TX, United States
[2]The University of Texas at Houston - Memorial Hermann Center for Healthcare Quality and Safety, Houston, TX, United States
[3]Department of Biostatistics, The University of Texas M.D. Anderson Cancer Center, Houston, TX, United States
[4]Department of Medicine, Penn State College of Medicine, Hershey, PA, United States

Corresponding Author:
Kevin O Hwang, MD, MPH
Department of Internal Medicine
The University of Texas Medical School at Houston
6410 Fannin St
UPB 1100.41
Houston, TX, 77030
United States
Phone: 1 (713) 500 6441
Fax: 1 (713) 500 0766
Email: kevin.o.hwang@uth.tmc.edu

Abstract

Background: Online weight loss programs are increasingly popular. However, little is known about outcomes and associations with website usage among members of free online weight loss programs.

Objective: This retrospective cohort study examined the association between website usage and weight loss among members of a free commercial online weight loss program (SparkPeople).

Methods: We conducted a retrospective analysis of a systematic random sample of members who joined the program during February 1 to April 30, 2008, and included follow-up data through May 10, 2010. The main outcome was net weight change based on self-reported weight. Measures of website usage included log-ins, self-monitoring entries (weight, food, exercise), and use of social support tools (discussion forums, friendships).

Results: The main sample included 1258 members with at least 2 weight entries. They were 90.7% female, with mean (SD) age 33.6 (11.0) and mean (SD) BMI 31.6 (7.7). Members with at least one forum post lost an additional 1.55 kg (95% CI 0.55 kg to 2.55 kg) relative to those with no forum posts. Having at least 4 log-in days, weight entry days, or food entry days per 30 days was significantly associated with weight loss. In the multiple regression analysis, members with at least 4 weight entry days per 30 days reported 5.09 kg (95% CI 3.29 kg to 6.88 kg) more weight loss per 30 days than those with fewer weight entry days. After controlling for weight entry days, the other website usage variables were not associated with weight change.

Conclusions: Weekly or more frequent self-monitoring of weight is associated with greater weight loss among members of this free online weight loss program.

(J Med Internet Res 2013;15(1):e11) doi:10.2196/jmir.2195

KEYWORDS

Internet; Obesity; Overweight; Weight loss; Adherence; Attrition

Introduction

More than one third of US adults are obese [1]. Online weight loss programs represent a convenient and scalable resource for the prevention and treatment of obesity. In order to maximize the efficacy of such programs, it is important to identify specific features that promote favorable weight loss outcomes [2].

A positive association between website usage (program engagement) and weight loss has been demonstrated among participants enrolled in randomized trials and other prospective investigations [3-11]. But because those studies involved research volunteers who met strict eligibility criteria, results may have limited applicability to the general public. Instead of enrolling in a formal research study, individuals seeking to lose weight are more likely to join a commercial online weight loss program. Naturalistic evaluations of commercial online weight loss programs are needed to understand the benefits and limitations of these increasingly popular resources. While recent reports have described outcomes among paid subscribers to commercial online weight loss programs based in Australia [12], the United Kingdom [13], and Sweden [14], comparatively little is known about website usage and weight loss among members of *free* online weight loss programs. This is an important gap in the literature because consumers and health care providers strongly prefer free weight loss programs [15,16].

Therefore, we evaluated a naturalistic cohort of members of SparkPeople, which is a free online weight loss program based in the United States. Prior studies described the accuracy of advice [17] and types of social support [18] shared among SparkPeople members, as well as the positive association between use of the program's online forums and perceived social support [19,20]. The purpose of this retrospective cohort study was to describe the magnitude of weight loss and examine the association between website usage and weight loss.

Methods

Membership in the SparkPeople online weight loss program is free and supported chiefly through advertising revenue. Most members are from the United States. The main features are educational content, self-monitoring tools (for weight, diet, and exercise), and social support venues (discussion forums, blogs, and "Friend" relationships similar to general online social networks). Members use the program largely in a self-directed and self-paced manner. They are free to use website features and make weight entries at any time. As of 2009, members could also access components of the website via mobile applications, although data on mobile access were not included in the present dataset.

Study Sample

Approximately 521,000 members joined the program during February 1 – April 30, 2008. Systematic random sampling produced the initial cohort of 26,582 individuals with a baseline weight. This de-identified dataset included all available follow-up data for these members through May 10, 2010.

Members were then excluded from analysis if they reported extreme outlying values for baseline characteristics: age greater than 100 years, weight less than 100 lb (45.4 kg) or greater than 800 lb (362.9 kg) with no follow-up weight entries, or height greater than 10 ft (3.05 m). These extreme outlier values were considered unrealistic, data entry errors, and/or not relevant to the analysis. Exclusion of 469 members with outlier values left an interim cohort of 26,113 members. Although we did not implement a filter for very short heights, the filtering process eliminated members with extreme height values, as the height range was 49-78 inches in the final cohort (n=1258) described below.

Self-reported weights may be inaccurate due to data entry errors, such as omitting, adding, or transposing digits. All 58,574 weight entries were analyzed for potential data entry errors with a series of automated and manual procedures. There were three automated filters (F1, F2, F3) for members with three or more weight entries, and a fourth automated filter (F4) for members with two or more weight entries. The first filter (F1) entailed fitting a second order polynomial regression line on the weight entries and dates; the distance of each point from the expected line (residual) was calculated. Points were flagged for visual evaluation if the residual was more than three times the standard deviation of the residuals for the individual record, and the difference between the actual and expected weight was greater than 10 lb (4.5 kg). The second filter (F2) flagged members that had a change of more than 2% body weight per day (whether over a short or long period of time). The third filter (F3) flagged members who had an absolute change of 50 lb (22.7 kg) or more between any two weight entries. The fourth filter (F4) flagged members with a weight change of 100 lb (45.4 kg) or more over the complete recorded period.

The automated filters identified 301 potentially erroneous weight entries among 248 unique SparkPeople members. Two independent observers (KOH and AWT) manually reviewed these 301 weight entries in the context of the other weight entries for a given individual (the trend) to determine whether the entry was erroneous. The observers demonstrated high interrater reliability, with an overall 95% agreement and Cohen's Kappa .90. When the observers disagreed about a weight entry, they discussed and reached a consensus determination. After the automated filters and manual review, 73 of 301 weight entries (24%) were deemed erroneous and excluded from analysis. This represents 0.12% of all weights available (73/58,574).

Because the main objective was to evaluate the relationship between website usage and weight loss, 20,518 members with only one weight entry were excluded. The subsequent interim cohort of 5595 members had at least two weight entries. Members with at least two weight entries were younger, with a mean (SD) age 35.3 (11.5) years vs. 37.2 (13.4) years, *P*<.001, heavier, with a mean (SD) BMI 31.6 (7.7) vs. 30.8 (20.2), *P*<.001, and more likely to be female (91.3% vs. 85.9%, *P*<.001) than those with only one weight entry.

Each weight entry and log-in event was date-stamped. Because the data did not include the dates for use of other website features (eg, exercise diary), it was possible that a member used a website feature after his or her last weight entry, thereby complicating the interpretation of associations between website usage and weight change. However, it was possible to determine when a member stopped all website usage activity because each website activity generated a date-stamped log-in. Therefore, the final cohort of n=1258 was defined as those who had at least two weight entries, with the last weight entry on the same day as or after utilization of other website features. In other words, this final cohort consists of members who used the website features between their first and last weight entries. While the study design does not allow definitive analysis of causation, the final cohort at least meets the temporality criteria for causation. This final cohort was analyzed to examine the relationships between website usage and weight change. The flow diagram (Figure 1) depicts how the final cohort for analysis was defined.

Figure 1. Flow diagram.

Definition of Website Usage Variables

Usage of website features included the following variables:

1. Log-in days: the number of days the member logged into the website at least once, regardless of whether s/he used any other website features while logged in.
2. Weight entry days: the number of days the member entered weight at least once.
3. Food entry days: the number of days the member made at least one entry in the food diary, regardless whether those entries were complete or incomplete.
4. Exercise entry days: the number of days the member made at least one entry in the exercise diary, regardless whether those entries were complete or incomplete.
5. Exercise minutes: the total number of minutes of exercise the member recorded in the exercise diary during the study period.
6. SparkPoints: the number of points earned by the member for miscellaneous website activities, such as reading articles, taking quizzes and polls, using food and exercise diaries, posting messages on forums, and making blog entries.
7. Forum posts: the number of messages the member posted on the discussion forums.

8. Friends: the number of other members designated as “SparkFriends” by the member (a social network feature).

Analysis of Weight Entry Span and Weight Change

The weight entry span was defined as the number of days between first and last weight entry. Weight change, defined as the last recorded weight minus baseline weight, was stratified by weight entry span.

General Analytic Approach to Website Usage and Weight Change

Potential confounders of the relationship between website usage and weight change included baseline BMI, age, gender, and weight entry span. Registration source (referred by friend, search engine ad, other) was examined as a potential confounder because it is possible that members who are introduced to the online program by a personal friend may differ in website usage and weight change than those who find the site only after an Internet search. Likewise, geographic location (zip code in midwest, northeast, south, or west region of the United States or other country) was examined as a potential confounder because of possible geographical variation in Internet use and dietary and physical activity factors related to body weight. These potential confounders were examined for association with weight change by using univariable regression analyses. Variables with a significant association ($P<.05$) with weight change were adjusted for in subsequent analyses.

Regression analysis was conducted with each website usage variable as predictor and net weight change as the outcome with adjustment for the identified confounders (univariable analyses). Multiple regression analysis with backward stepwise selection was conducted to identify the best subset of predictive covariates for net weight change. All covariates were included in the full model before model selection. Variables were included in the final model if $P<.20$.

Analysis of Binary Website Usage Variables and Weight Change

The distribution of website usage was highly skewed, such that many members did not use a given feature before their last weight entry. Therefore, website usage variables were coded as binary variables for the initial analysis (*never* using a website feature vs. using it *at least once*). Binary coding of website usage has been used in a prior study of an online weight loss program [21]. Weight change was defined as the last recorded weight minus baseline weight. Because all members in the sample had two or more log-ins and weigh-ins, and most members had at least one SparkPoint, those variables were not included in this analysis.

Analysis of Categorical Website Usage Variables and Weight Change

To further analyze associations between website usage and weight change, one option would be to assume a linear regression model, but the linearity assumption may be not be valid. For example, the change in weight associated with an increase from 3 to 4 forum posts may be different from the weight change associated with an increase from 150 to 151 forum posts. We considered categorizing the website usage variables into tertiles, but the data distributions were too skewed, so that observations in the top tertile may be equivalent to some observations in the middle tertile.

Weekly use of online weight loss program features has been associated with greater perceived social support [20] and weight loss [3,4,6,7,14]. Therefore, we categorized log-in days, weight entry days, food entry days, exercise entry days, and forum posts as greater than or less than 4 events per 30 days, corresponding to approximately once per week. The median value of SparkPoints was 20.8 per 30 days, so SparkPoints was categorized as above or below 20 per 30 days. We initially categorized exercise minutes as above or below 600 minutes per 30 days, corresponding approximately to 150 minutes per week, but there were not enough members who met this threshold to allow meaningful analysis. Therefore, we categorized exercise minutes as above or below 120 minutes per 30 days, corresponding to 4 exercise entries of 30 minutes length per 30 days. Zero use was also included as a category for all variables except log-in days and weight entry days because, by definition, all members in the analysis had at least one log-in day and weight entry day.

Weight change was defined as (last recorded weight minus baseline weight) divided by 30 day unit of weight entry span. For example, if a member’s last weight was 100 kg and baseline weight was 108 kg, and weight entry span was 60 days, then the weight change was $-8 / 2 = -4$ kg per 30 days.

Matlab version R2008b (MathWorks, Natick, MA) was used to identify potential outlying values for baseline characteristics and self-reported weights. The R statistical package (R Development Core Team, Version 2.14.1) was used for descriptive and regression analyses.

The study was approved by the Committee for the Protection of Human Subjects at the University of Texas Health Science Center at Houston.

Results

Demographic characteristics of the final cohort are summarized in Table 1.

Most members did not use the website features frequently. Since values were skewed, they are summarized with medians and interquartile ranges in Table 2.

Table 1. Demographic characteristics of final cohort (n=1258).

Characteristics		mean (SD) or n (%)
Age, years, mean (SD)		33.6 (11.0)
Gender, female, n (%)		1141 (90.7)
Baseline BMI, mean (SD)		31.6 (7.7)
Geographic location, n (%)		
	US - Midwest	291 (23.1)
	US - Northeast	187 (14.9)
	US - South	400 (31.8)
	US - West	210 (16.7)
	Other country	170 (13.5)

Table 2. Website usage characteristics of final cohort (n=1258).

	Minimum	25th percentile	Median	75th percentile	Maximum
Log-in days	2	3	7	23	654
Log-in span (number of days between first and last log-in days)	1	54	306.5	624.8	826
Weight entry days	2	2	2	4	156
Weight entry span (number of days between first and last weight entry days)	1	54	306	623.8	826
Food entry days	0	0	2	9	270
Exercise entry days	0	0	1	4	258
Exercise minutes	0	0	40	256.5	39130
Friends	0	0	0	1	74
Forum posts	0	0	0	0	2758
SparkPoints	0	22.3	85	330	16480

Table 3. Total weight change stratified by weight entry span[a] among final cohort (n=1258).

Weight entry span	n (%)	Mean (SD) total weight change
Up to 30 days	261 (20.7%)	-1.08kg (3.09)
31-60 days	69 (5.5%)	-2.00kg (2.68)
61-90 days	39 (3.1%)	-1.98kg (3.46)
91-183 days	131 (10.4%)	-2.32kg (6.50)
184-365 days	210 (16.7%)	-1.97kg (9.27)
366-548 days	178 (14.1%)	-0.99kg (8.04)
549-829 days	370 (29.4%)	-1.38kg (10.5)

[a] Weight entry span is the number of days between first and last weight entry.

Table 3 shows the total weight change stratified by weight entry span.

Baseline BMI was the only potential confounder significantly associated with weight change. Each additional unit of BMI at baseline was associated with an additional reported weight loss of 0.28 kg (95% CI 0.22 to 0.33 kg, *P*<.001) during the weight entry span. Age, gender, referral source, geographic location, and weight entry span were not associated with weight change. Further analyses of predictors of weight change are adjusted for baseline BMI.

Website Usage as Binary Variables

In the analysis with binary website usage variables, having at least one forum post was the only website variable significantly associated with greater weight loss. Members with at least one forum post reported an additional weight loss of 1.55 kg (95%

CI 0.55 kg to 2.55 kg) relative to those with no forum posts (Table 4).

Table 4. Associations between website usage variables (binary) and weight change[a] in the final cohort (n=1258).

	N (%)	Coefficient	95% CI	*P* value
At least one food entry day	893 (71.0)	-0.35 kg	-1.28 kg to 0.58 kg	.47
At least one exercise entry day	677 (53.8)	-0.67 kg	-1.52 kg to 0.18 kg	.12
At least one exercise minute	741 (58.9)	-0.61 kg	-1.47 kg to 0.25 kg	.16
At least one forum post	291 (23.1)	-1.55 kg	-2.55 kg to -0.55 kg	.002
At least one Friend	547 (43.5)	-0.80 kg	-1.66 kg to 0.06 kg	.07

[a] Website usage variables were coded as binary (at least once versus never). Analyses were adjusted for baseline BMI. Because all members in the sample had two or more log-ins and weigh-ins, and most members had at least one SparkPoint, those variables were not included.

Website Usage as Categorical Variables

In the analyses with categorical website usage variables, weekly log-ins, weight entries, and food entries (at least 4 in 30 days) were associated with greater reported weight loss than less frequent use (Table 5). More exercise entry days per 30 days was also positively associated with weight loss. SparkPoints was associated with relative weight gain compared to having zero SparkPoints (although only 74 of 1258 members had zero SparkPoints). On multiple regression analysis including baseline BMI and all website usage variables, weight entry days was the only variable significantly associated with weight change. Members with at least 4 weight entry days per 30 days (weekly) reported an additional 5.09 kg weight loss per 30 days (95% CI 3.29 kg to 6.88 kg) relative to those with fewer weight entry days.

Table 5. Univariable analyses of associations between website usage variables (categorized) per 30 days and reported weight change per 30 days[a] in the final cohort (n=1258).

		Sample size	Coefficient	95% CI	*P* value
Log-in days per 30 days					
	0 to <4	842	reference		Overall: <.001
	≥4	416	-3.18 kg	-4.69 kg to -1.67 kg	<.001
Weight entry days per 30 days					
	0 to <4	1018	reference		Overall: <.001
	≥4	240	-5.09 kg	-6.89 kg to -3.29 kg	<.001
Food entry days per 30 days					
	0	365	reference		Overall:.002
	1 to <4	684	1.04 kg	-0.59 kg to 2.68 kg	.21
	≥4	209	-2.49 kg	-4.68 kg to -0.30 kg	.03
Exercise entry days per 30 days					
	0	581	reference		Overall: .021
	1 to <4	576	1.18 kg	-0.30 kg to 2.67 kg	.12
	≥4	101	-2.46 kg	-5.20 kg to 0.26 kg	.08
Exercise minutes per 30 days					
	0	517	reference		Overall: .054
	1 to <120	556	1.43kg	-0.12 kg to 2.98 kg	.07
	≥ 120	185	-0.89kg	-3.06 kg to 1.28 kg	.42
SparkPoints per 30 days					
	0	74	reference		Overall: .003
	1 to <20	546	5.35 kg	2.22 kg to 8.48 kg	<.001
	≥20	638	4.47 kg	1.37 kg to 7.58 kg	.005
Forum posts per 30 days					
	0	967	reference		Overall: .24
	1 to <4	233	0.99 kg	-0.86 kg to 2.84 kg	.30
	≥4	58	-2.09 kg	-5.52 kg to 1.33 kg	.23
Friends					
	0	711	reference		Overall: .55
	1	327	0.88kg	-0.81 kg to 2.57 kg	.31
	≥2	220	0.90 kg	-1.08 kg to 2.87 kg	.37

[a] Analyses were adjusted for baseline BMI.

Discussion

Key Findings

Average weight loss, based on self-report, was modest in this free online weight loss program, but active users had better outcomes. Making weekly weight entries (at least 4 weight entry days per 30 days) was associated with an additional 5 kg weight loss. After controlling for weight entry days, the other website usage variables were not associated with weight loss.

Comparison With Prior Studies

To our knowledge, this is the first analysis of a naturalistic cohort of members of a free online weight loss program available to the general public. Prior studies evaluated naturalistic cohorts of members who paid for monthly subscriptions to commercial online weight loss programs in Australia [12], the UK [13], and Sweden [14]. Because free programs may have greater potential to reach people in need of weight loss assistance [15,16], evaluating this free program is an important advancement in our understanding of weight loss resources. Another strength of this study was the use of multiple regression to identify variables that are independently associated with weight loss.

Our results are consistent with prior studies documenting a positive relationship between engagement in online programs and weight control. Most of these involved research volunteers and strict eligibility criteria [3-11], while others were naturalistic studies of commercial online programs [12-14]. Overall, favorable weight outcomes were associated with higher frequencies of log-ins [3-7,9,10,12,14], weight entries [6,10,11,14], self-monitoring entries for diet and/or exercise [6-8,10-13], and use of social support tools [6,8,12,13].

Although the current study discovered that several website usage variables were associated with weight loss, the multiple regression analysis indicates that weight entry days is the most important. It is possible that weight self-monitoring leads to frequent modification of diet and exercise behavior in response to weights. An alternative explanation is that members made weight entries only when they were losing (rather than gaining) weight. However, a post-hoc analysis found that 71% percent of members posted at least one weight reflecting a higher weight than a previous entry. In other words, they did not avoid documenting weight regain. The evidence in favor of weight self-monitoring improving weight control is strengthened by the consistency of the association in studies of online [6,10,11,14] and traditional [22-28] programs. When the analysis controlled for weight entry days, the other website usage variables were not related to weight change. It may be that the main benefit of these other website features is to maintain interest in the online program and maximize the opportunity for weight self-monitoring.

When website usage variables were coded as binary variables, making at least one forum post was associated with greater weight loss, although this analysis did not adjust for weight entry days. Because the content of the messages was not available for this study, we could not determine the nature of interactions on the forums. Furthermore, it was not possible to assess how often members read messages on a forum without posting ("lurking"). However, prior studies found that SparkPeople members receive high-quality advice [17] and social support [18] on the forums and that use of forums is associated with greater perceptions of social support [19,20]. Forum use may facilitate weight loss in part by boosting adherence to weight self-monitoring. Although the optimal source, frequency, and venue for social support has not been determined, the link between social support and weight loss has been reported in several studies of online [6,8,12,13] and traditional [29-34] weight loss programs.

Limitations

The study had several limitations. First, attrition was high, which is a common problem plaguing commercial online weight loss programs [12-14]. The majority of members in the initial sample made only one weight entry, so that the main analytic sample represented a small portion of all who registered for the program. Compared to those with only one weight entry, members with multiple weight entries were younger, heavier, and more likely to be female—consistent with the profile of individuals who typically enroll in weight loss programs. Even within the analytic sample, 21% of members quit entering their weight after 30 days. Since it is very easy to join a *free* online program, it is possible that many people joined without a strong sense of commitment, thereby inflating the initial sample and the subsequent attrition rate. Sending personalized follow-up messages to inactive members may improve adherence to weight self-monitoring [7], although this approach would be difficult to implement without additional resources.

Another limitation of the study is that the available records did not include ethnicity, race, or clinical characteristics. Furthermore, members were predominantly female, as is typical of weight loss programs. Taken together, these limitations indicate that results cannot be extrapolated to the general population, regular Internet users, or those who initially who register for this online program. The study is also limited by reliance on self-reported weights. However, weight reported by members of another online weight loss program was highly correlated with and similar to objective weights [35]. The lack of date stamps on most website usage variables (except log-ins and weight entries) limited the depth of analysis on the intensity and patterns of website usage. Lastly, because this was an observational study, the causal link between usage of website features and weight loss cannot be definitively ascertained without experimental studies.

Conclusions

The public health impact of an intervention is determined by efficacy and dissemination [36]. Because this online program is free, scalable, and widely disseminated, the potential public health impact is significant. The current study suggests that increasing the use of weight self-monitoring among current and future members may boost the impact of this program even more.

Acknowledgments

We thank David Heilmann and Josh Knepfle at SparkPeople for providing the dataset. We also thank Weiwei Wang, PhD, for assisting with preliminary data analyses.

This work was supported in part by a Clinical Investigator Award from The University of Texas Health Science Center at Houston, Center for Clinical Research and Evidence-Based Medicine and The University of Texas at Houston - Memorial Hermann Center for Healthcare Quality and Safety. The funding sources were not involved in reviewing or approving the manuscript for publication.

Conflicts of Interest

None declared.

References

1. Flegal KM, Carroll MD, Kit BK, Ogden CL. Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999-2010. JAMA 2012 Feb 1;307(5):491-497. [doi: 10.1001/jama.2012.39] [Medline: 22253363]
2. 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]
3. 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]
4. 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]
5. 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]
6. Gold BC, Burke S, Pintauro S, Buzzell P, Harvey-Berino J. Weight loss on the web: A pilot study comparing a structured behavioral intervention to a commercial program. Obesity (Silver Spring) 2007 Jan;15(1):155-164 [FREE Full text] [doi: 10.1038/oby.2007.520] [Medline: 17228043]
7. 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;166(15):1620-1625. [doi: 10.1001/archinte.166.15.1620] [Medline: 16908795]
8. Krukowski RA, Harvey-Berino J, Ashikaga T, Thomas CS, Micco N. Internet-based weight control: the relationship between web features and weight loss. Telemed J E Health 2008 Oct;14(8):775-782 [FREE Full text] [doi: 10.1089/tmj.2007.0132] [Medline: 18954247]
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. 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]
11. 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 [FREE Full text] [doi: 10.1038/oby.2010.119] [Medline: 20523304]
12. Neve M, Morgan PJ, Collins CE. Weight change in a commercial web-based weight loss program and its association with website use: cohort study. J Med Internet Res 2011;13(4):e83 [FREE Full text] [doi: 10.2196/jmir.1756] [Medline: 21993231]
13. Johnson F, Wardle J. The association between weight loss and engagement with a web-based food and exercise diary in a commercial weight loss programme: a retrospective analysis. Int J Behav Nutr Phys Act 2011;8:83 [FREE Full text] [doi: 10.1186/1479-5868-8-83] [Medline: 21810222]
14. Jonasson J, Linné Y, Neovius M, Rössner S. An Internet-based weight loss programme -- a feasibility study with preliminary results from 4209 completers. Scand J Public Health 2009 Jan;37(1):75-82. [doi: 10.1177/1403494807085066] [Medline: 18684784]
15. Krist AH, Woolf SH, Johnson RE, Rothemich SF, Cunningham TD, Jones RM, et al. Patient costs as a barrier to intensive health behavior counseling. Am J Prev Med 2010 Mar;38(3):344-348. [doi: 10.1016/j.amepre.2009.11.010] [Medline: 20171538]
16. Hwang KO, Stuckey HL, Chen MC, Kraschnewski JL, Forjuoh SN, Poger JM, et al. Primary care providers' perspectives on online weight-loss programs: a big wish list. J Med Internet Res 2012;14(1):e16 [FREE Full text] [doi: 10.2196/jmir.1955] [Medline: 22262726]
17. Hwang KO, Farheen K, Johnson CW, Thomas EJ, Barnes AS, Bernstam EV. Quality of weight loss advice on internet forums. Am J Med 2007 Jul;120(7):604-609 [FREE Full text] [doi: 10.1016/j.amjmed.2007.04.017] [Medline: 17602934]
18. Hwang KO, Ottenbacher AJ, Green AP, Cannon-Diehl MR, Richardson O, Bernstam EV, et al. Social support in an Internet weight loss community. Int J Med Inform 2010 Jan;79(1):5-13 [FREE Full text] [doi: 10.1016/j.ijmedinf.2009.10.003] [Medline: 19945338]
19. Hwang KO, Ottenbacher AJ, Lucke JF, Etchegaray JM, Graham AL, Thomas EJ, et al. Measuring social support for weight loss in an internet weight loss community. J Health Commun 2011 Feb;16(2):198-211. [doi: 10.1080/10810730.2010.535106] [Medline: 21181600]
20. Hwang KO, Etchegaray JM, Sciamanna CN, Bernstam EV, Thomas EJ. Structural social support predicts functional social support in an online weight loss programme. Health Expect 2012 Jan 2. [doi: 10.1111/j.1369-7625.2011.00759.x] [Medline: 22212418]
21. 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]

22. Vanwormer JJ, French SA, Pereira MA, Welsh EM. The impact of regular self-weighing on weight management: a systematic literature review. Int J Behav Nutr Phys Act 2008;5:54 [FREE Full text] [doi: 10.1186/1479-5868-5-54] [Medline: 18983667]
23. 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]
24. Linde JA, Jeffery RW, French SA, Pronk NP, Boyle RG. Self-weighing in weight gain prevention and weight loss trials. Ann Behav Med 2005 Dec;30(3):210-216. [doi: 10.1207/s15324796abm3003_5] [Medline: 16336072]
25. Wing RR, Tate DF, Gorin AA, Raynor HA, Fava JL. A self-regulation program for maintenance of weight loss. N Engl J Med 2006 Oct 12;355(15):1563-1571. [doi: 10.1056/NEJMoa061883] [Medline: 17035649]
26. Butryn ML, Phelan S, Hill JO, Wing RR. Consistent self-monitoring of weight: a key component of successful weight loss maintenance. Obesity (Silver Spring) 2007 Dec;15(12):3091-3096 [FREE Full text] [doi: 10.1038/oby.2007.368] [Medline: 18198319]
27. VanWormer JJ, Martinez AM, Martinson BC, Crain AL, Benson GA, Cosentino DL, et al. Self-weighing promotes weight loss for obese adults. Am J Prev Med 2009 Jan;36(1):70-73. [doi: 10.1016/j.amepre.2008.09.022] [Medline: 18976879]
28. Gokee-Larose J, Gorin AA, Wing RR. Behavioral self-regulation for weight loss in young adults: a randomized controlled trial. Int J Behav Nutr Phys Act 2009;6:10 [FREE Full text] [doi: 10.1186/1479-5868-6-10] [Medline: 19220909]
29. 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]
30. Elfhag K, Rössner S. Who succeeds in maintaining weight loss? A conceptual review of factors associated with weight loss maintenance and weight regain. Obes Rev 2005 Feb;6(1):67-85. [doi: 10.1111/j.1467-789X.2005.00170.x] [Medline: 15655039]
31. Verheijden MW, Bakx JC, van Weel C, Koelen MA, van Staveren WA. Role of social support in lifestyle-focused weight management interventions. Eur J Clin Nutr 2005 Aug;59 Suppl 1:S179-S186. [doi: 10.1038/sj.ejcn.1602194] [Medline: 16052189]
32. Gorin A, Phelan S, Tate D, Sherwood N, Jeffery R, Wing R. Involving support partners in obesity treatment. J Consult Clin Psychol 2005 Apr;73(2):341-343. [doi: 10.1037/0022-006X.73.2.341] [Medline: 15796642]
33. Gallagher KI, Jakicic JM, Napolitano MA, Marcus BH. Psychosocial factors related to physical activity and weight loss in overweight women. Med Sci Sports Exerc 2006 May;38(5):971-980. [doi: 10.1249/01.mss.0000218137.25970.c6] [Medline: 16672853]
34. Kumanyika SK, Wadden TA, Shults J, Fassbender JE, Brown SD, Bowman MA, et al. Trial of family and friend support for weight loss in African American adults. Arch Intern Med 2009 Oct 26;169(19):1795-1804. [doi: 10.1001/archinternmed.2009.337] [Medline: 19858438]
35. Harvey-Berino J, Krukowski RA, Buzzell P, Ogden D, Skelly J, West DS. The accuracy of weight reported in a web-based obesity treatment program. Telemed J E Health 2011 Nov;17(9):696-699 [FREE Full text] [doi: 10.1089/tmj.2011.0032] [Medline: 21882997]
36. 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]

Edited by G Eysenbach; submitted 31.05.12; peer-reviewed by L Ware, M Waring; comments to author 23.06.12; revised version received 29.08.12; accepted 08.11.12; published 15.01.13.

Please cite as:
Hwang KO, Ning J, Trickey AW, Sciamanna CN
Website Usage and Weight Loss in a Free Commercial Online Weight Loss Program: Retrospective Cohort Study
J Med Internet Res 2013;15(1):e11
URL: http://www.jmir.org/2013/1/e11/
doi:10.2196/jmir.2195
PMID:23322819

Original Paper

An Interactive Internet-Based Plate for Assessing Lunchtime Food Intake: A Validation Study on Male Employees

Madeleine Svensson[1,2], PhD; Rino Bellocco[3,4], ScD; Linda Bakkman[1], PhD; Ylva Trolle Lagerros[1], PhD, MD

[1]Karolinska Institutet, Department of Medicine, Unit of Clinical Epidemiology, Stockholm, Sweden
[2]School of Health and Social Sciences, Halmstad University, Halmstad, Sweden
[3]Karolinska Institutet, Department of Medical Epidemiology and Biostatistics, Stockholm, Sweden
[4]University of Milano-Bicocca, Department of Statistics, Milan, Italy

Corresponding Author:
Madeleine Svensson, PhD
Karolinska Institutet
Department of Medicine
Unit of Clinical Epidemiology
T2
Stockholm, 171 76
Sweden
Phone: 46 8 517 791 73
Fax: 46 8 517 793 04
Email: Madeleine.Svensson@hh.se

Abstract

Background: Misreporting food intake is common because most health screenings rely on self-reports. The more accurate methods (eg, weighing food) are costly, time consuming, and impractical.

Objectives: We developed a new instrument for reporting food intake—an Internet-based interactive virtual food plate. The objective of this study was to validate this instrument's ability to assess lunch intake.

Methods: Participants were asked to compose an ordinary lunch meal using both a virtual and a real lunch plate (with real food on a real plate). The participants ate their real lunch meals on-site. Before and after pictures of the composed lunch meals were taken. Both meals included identical food items. Participants were randomized to start with either instrument. The 2 instruments were compared using correlation and concordance measures (total energy intake, nutritional components, quantity of food, and participant characteristics).

Results: A total of 55 men (median age: 45 years, median body mass index [BMI]: 25.8 kg/m^2) participated. We found an overall overestimation of reported median energy intake using the computer plate (3044 kJ, interquartile range [IQR] 1202 kJ) compared with the real lunch plate (2734 kJ, IQR 1051 kJ, $P<.001$). Spearman rank correlations and concordance correlations for energy intake and nutritional components ranged between 0.58 to 0.79 and 0.65 to 0.81, respectively.

Conclusion: Although it slightly overestimated, our computer plate provides promising results in assessing lunch intake.

(J Med Internet Res 2013;15(1):e13) doi:10.2196/jmir.2217

KEYWORDS
diet; epidemiology; Internet; methods; nutrition; validation; Web

Introduction

Measuring food intake is a challenge. Most assessment tools rely on an individual's ability to accurately recall and report foods consumed, usually according to a fixed format of an instrument [1]. Examples of traditional methods to examine food intake include food frequency questionnaires (FFQ), 24-hour recalls, and food recording and weighing [1-3], as well as the duplicate-portion technique [4]. Weighing individuals' food plates with individually composed meals before and after eating is the most precise method, but it can be a rather costly, time-consuming, and impractical approach. Hence, self-reported food intake is typically used in health screenings.

One of the challenges of self-reported food intake is the high rate of misreporting [5,6]. Overweight or obese women and individuals of low socioeconomic status [7] tend to underreport food consumption. Food items that are sweet, fatty, and considered unhealthy are more likely to be underreported. In contrast, food with high protein content or vegetables and fruits are frequently exaggerated [3,7-12]. Although some validation studies present accurate measurements of food intake [1], respondents' may still struggle with reporting food intake because of extensive questionnaires that are difficult to fill out [8].

A recent study conducted by Illner et al [13] reports a similar degree of misreporting of food intake irrespective of method of delivery. More specifically, the participants' food intake reporting was identical using paper-based frequency assessments and technology-based assessments (ie, Internet-based). Yet, the benefits gained from using technology in food assessments may speak for an increased interest and usage in nutritional research [13,14] compared with conventional methods. The Internet promotes time- and cost-effective research and facilitates administration of research material, as well as collection and storage of data [14]. In addition, it allows for interactivity that, in turn, produces opportunities for the development of pedagogical advancements [15].

Pictures of foods and meal compositions have been used to facilitate reporting of food intake in prior nutritional research [16,17]. For instance, Turconi et al [16] asked their study participants to estimate food intake by looking at pictures of prepared meals of different portion sizes (small, medium, and large) put together in food atlases. The estimated meals were then compared with the participants' intake of actual meals, and indicated promising results on the participants' overall comprehension of food intake. Elinder et al [18] also reported valid results from allowing individuals with intellectual disabilities to photograph meals before and after intake. Hence, the use of pictures seems as an appropriate strategy in food intake assessments.

To incorporate the advantages of technology and visuals in nutritional research, we developed an Internet-based virtual food plate to measure lunch intake using the computer. Our computerized food plate allows for interactive composition of a single lunch meal, in which the user can add or subtract pictures of food items onto a virtual plate. To our knowledge, this format of food intake assessments has not been described previously. Consequently, the present study aimed to validate our new instrument against the golden standard—the participants' real lunch meal composition using real food items and utensils.

Methods

Participants

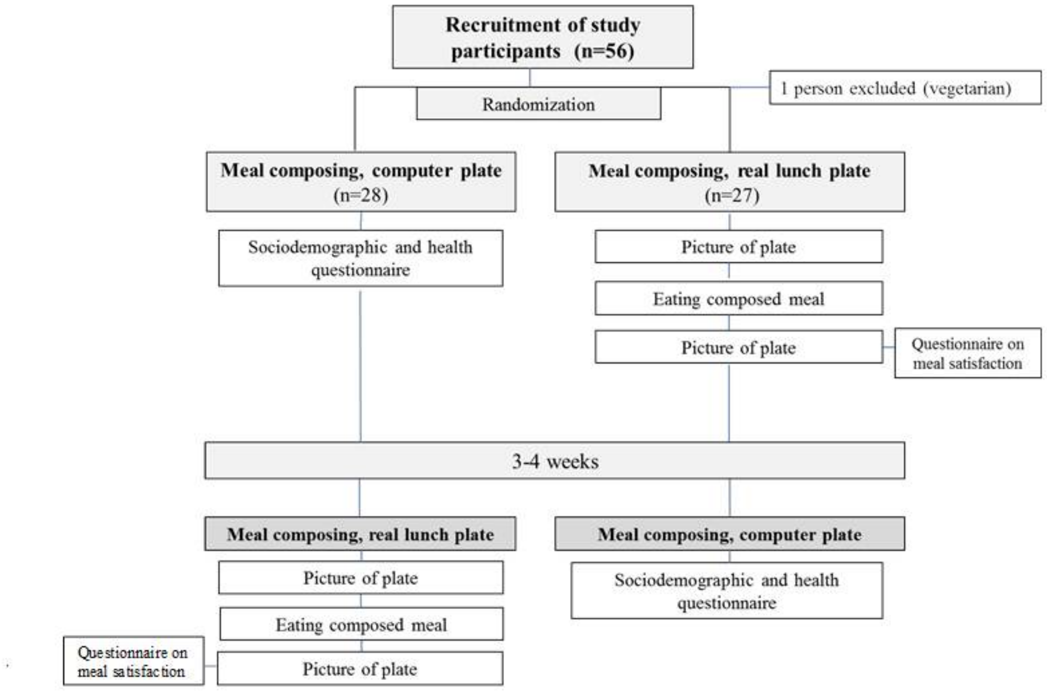
Between February and April 2010, 56 male employees (age 18-65 years) at the Swedish Transport Administration, Stockholm, Sweden, were asked to participate in the study. The predetermined food items of our instrument did not include vegetarian protein sources; therefore, one potential study participant who reported being a vegetarian was excluded from participating. Hence, a total of 55 employees participated in the present study. Most of the participants were employed as engineers (ie, work in an office).

Study Design

This validation study included two parts, using identical food items. The participants were asked to compose a lunch meal that represented their usual intake by means of (1) an interactive Internet-based food plate by adding suggested food items onto a virtual plate on the computer (hereafter referred to as "computer plate"), and (2) an ordinary lunch plate by adding real food items onto a real plate during a lunch setting (hereafter referred to as "real lunch plate").

The participants were recruited by the researchers in the company's main lobby during lunch hours (10 am - 1 pm). Upon recruitment, 28 participants were asked to start with the computer plate and 28 participants were asked to start with the real lunch plate. They were instructed to complete the remaining part (the computer plate or the real lunch plate) after 3 to 4 weeks. See Figure 1 for a flowchart of the study design. All participants signed an informed consent form prior to study start. The study was approved by the Karolinska Institutet's Ethical Committee in Stockholm, Sweden.

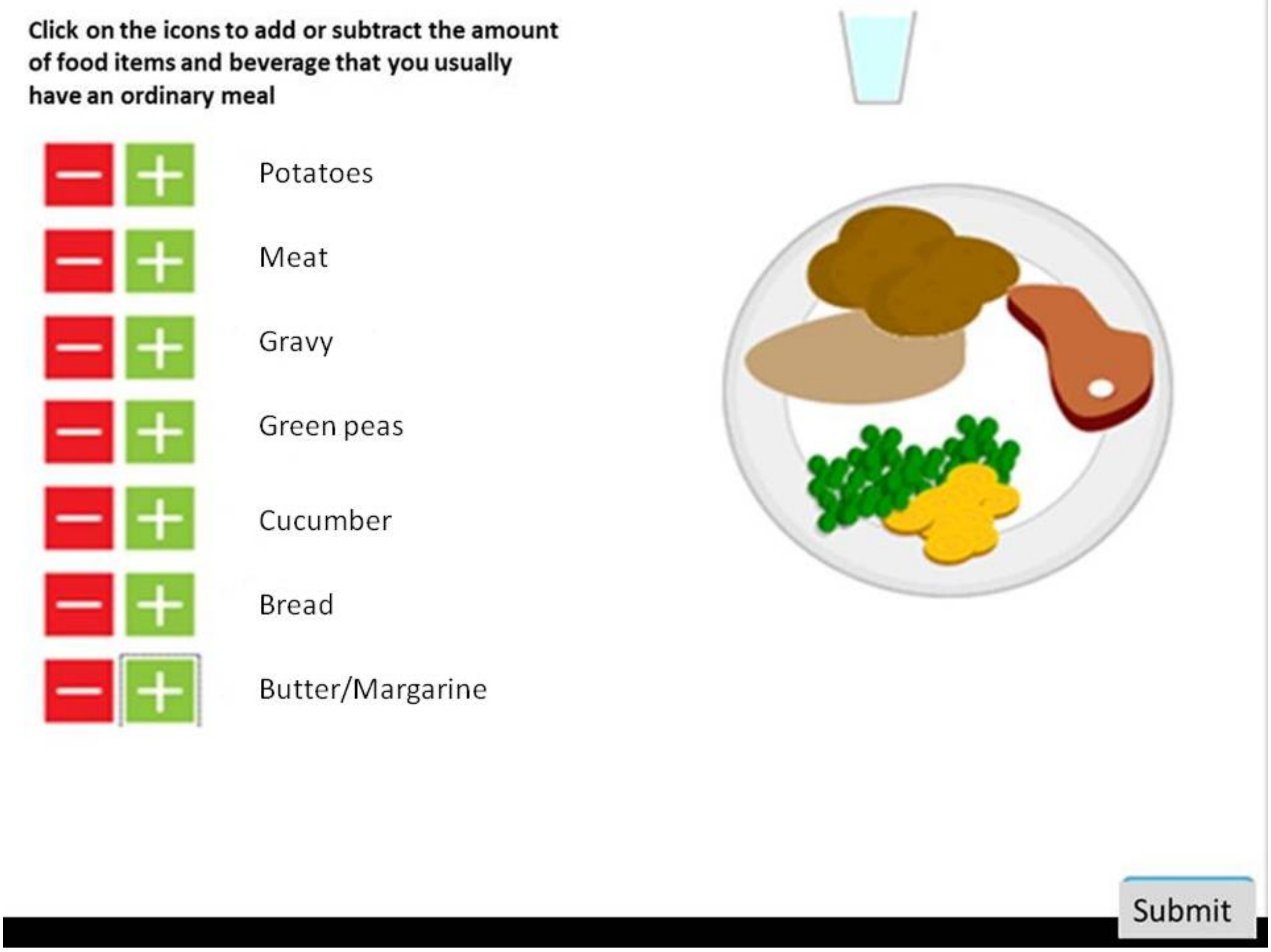
Figure 1. Flowchart of the validation process.



The Computer Plate

We developed a computer plate, an interactive instrument designed as a virtual food plate, available on a website. The virtual plate was placed in the center of the Web page, with a list of food items to the left. By clicking on the “+” and “–” buttons with the computer mouse, computer-generated pictures of food items were added to or subtracted from the virtual plate. It was possible to increase the quantity (or vice versa) of a food item by clicking several times. The participants had 7 food items to choose from when composing their meal on the computer, including boiled potatoes, meat (pork chops), gravy, green peas, slices of cucumber, slices of bread, and butter. Five different beverages were offered, including light beer (<3.5% alcohol), strong beer (≥ 3.5% alcohol), juice, milk (1.5% fat), and water. The food items were chosen because they are commonly represented in a Swedish lunch meal. See Figure 2 for an illustration of a composed lunch meal using the computer plate. (All beverages are not visible in Figure 2, but they appeared in the upper right corner of the website upon completion of the lunch meal assessment).

Figure 2. Example of the computer plate and listed food items.



Click on the icons to add or subtract the amount of food items and beverage that you usually have an ordinary meal

Study Part 1: The Computer Plate

The participants arrived at the main lobby and were instructed on how to complete the computer plate on a laptop provided by the researchers. The participants were informed that they could only build one plate representing a usual lunch meal, meaning that if they usually refilled their lunch plates with food, this refill had to be considered when composing the virtual meal. When the meal was composed, we saved the screen picture of the lunch meal in a Word document.

Next, the participants completed a questionnaire on sociodemographics, body mass index (BMI), physical activity level (PAL) [19], smoking habits, food allergies, and intake of breakfast or snacks prior to composing their computerized lunch meal.

Study Part 2: The Real Lunch Plate

The participants arrived to their workplace lunch area at lunchtime. Again, the participants were informed that they could only build one plate of a usual lunch meal (if they typically refilled their lunch plates, this amount of food had to be considered when composing the real lunch meal). The participants ate their lunch meal in the lunch area. Photos of the composed meals were taken prior and subsequent to eating.

After the participants composed their lunches, they were asked to fill out a questionnaire (described previously). In addition, they were asked to rate their level of fullness on a scale of 1 to 10 [20] to examine whether their reported lunch intake represented an appropriate meal intake.

The meals were not weighed because the purpose of our study was to examine whether the participants were able to visualize an ordinary lunch meal by using the pictures of food items provided in our computer plate instrument. We focused on the quantity of food items and overall nutritional content rather than the weight of food.

Statistical Analyses

We excluded 4 participants because of incomplete data. Descriptive statistics, such as median and interquartile range (IQR), were computed to summarize the participants' characteristics (eg, age, BMI, PAL, and meal satisfaction). Further, we computed summary statistics of the participants' composed meals with respect to total energy intake in kilojoules (kJ), as well as quantity of food items and nutritional components, including total energy intake (EI) and energy percentage (E%) of carbohydrates, proteins, and fats of the 2 meals. The before and after pictures of the composed meals were used to calculate the participants' quantities of food items included in the lunch meal.

We focused our analyses on total EI from food items on the plate excluding energy from beverages and bread/butter. The reason for this focus was because of the large energy differences among the various beverages to choose from (ie, the low/zero energy from water versus the high energy from beer). Wilcoxon signed rank tests and Wilcoxon rank sum tests [21] were used to test the difference between medians of the computer plate and real lunch plate measurements, and to assess if such differences were dependent on variables and/or on specific participant characteristics.

Participants were categorized into two groups based on the order of completing the 2 meal assessments: (1) computer plate-real lunch plate or (2) real lunch plate-computer plate. The participants were also categorized according to their BMI as normal weight (< 25 kg/m^2) or overweight (≥ 25 kg/m^2). Age was categorized as < 45 years or ≥ 45 years to study any differences in these characteristics with respect to meal composition. The cutoff age of 45 years was chosen because the median age was 45 years.

We used the Spearman rank correlation (ρ) to study the linear association among the reported EI, nutritional components, and food items using the computer plate and real lunch plate. The Spearman rank correlation is a nonparametric test that ranks 2 sets of outcomes distinctly and calculates a coefficient of rank correlation [22]. To further study the associations, we used the concordance correlation coefficient (ρ_c), an intraclass correlation that duplicates readings as replicates (random) rather than two distinct readings. It “evaluates the agreement between the two readings by measuring the variation from a 45° line through the origin (degree of concordance)” [23]. In addition, Bland-Altman plots were used to assess the differences between means of EI (kJ) estimated from the computer plate and the real lunch plate, plotted against the mean energy intake from the 2 methods [24]. To interpret the agreement, we considered participants with reported EI within the interval of ±10% from the mean (3014 kJ) of the real lunch plate as acceptable values of our new instrument.

Correlations and 95% confidence intervals (CI) were computed for all participants and stratified by groups (defined previously). All statistics were computed using the real lunch plate as the golden standard (reference). A significance level of .05 was used. Stata version 12 (Statacorp LP, College Station, TX, USA) was used for all statistical calculations and analyses.

Results

A total of 51 employees participated in the study. The participants had a median age of 45 years (IQR 21 years), a median BMI of 25.8 kg/m^2 (IQR 4.18 kg/m^2), and a median physical activity level of 1.65 PAL (IQR 0.1 PAL) (Table 1).

Table 1. Descriptive statistics of the study participants (N=51).

Description	Participants
Sex (males), n (%)	51 (100)
Age (years), median (IQR)	45 (21)
Body mass index (kg/m^2), median (IQR)	25.8 (4.18)
Physical activity level, median (IQR)	1.65 (0.1)
Smoker, n (%)	2 (3.7)
Had breakfast before study participation, n (%)	48 (87)

The participants’ reported total EI was somewhat higher on the computer plate compared to the real lunch meal. Overall, the median reported EIs for the computer plate and real lunch plate were 3044 kJ (IQR 1202 kJ) and 2734 kJ (IQR 1051 kJ, $P<.001$), respectively (Table 2). Although not significantly different, we noted that the participants with a BMI ≥ 25 kg/m^2 (Table 3) and those aged ≥ 45 years (data not shown) reported lower overestimations of EI (+147 kJ and +193 kJ, respectively) using the computer plate, compared to their counterparts (BMI < 25 kg/m^2: +595 kJ, $P=.90$; age < 45 years: +649 kJ, $P=.33$). Also, the EI measured from the computer plate was 172 kJ higher for participants starting with the real lunch plate compared to those starting with the computer plate (+729 kJ), although not statistically significant ($P=.75$).

Table 2. Study participants' composed meals using the two meal instruments (N=51).

Composed meal	Median (IQR)	
	Computer plate	Lunch plate
Reported intake [a]		
Total energy, kJ	3044 (1202)	2734 (1051)
Total energy including drinks, kJ	3341 (1348)	2989 (1277)
Total food, g	855 (189)	779 (248)
Total carbohydrates, E%	128 (57)	134 (67)
Total carbohydrates, g	59 (24)	55 (18)
Total proteins, E%	121 (29)	121 (29)
Total proteins, g	43 (32)	41 (31)
Total fat, E%	167 (29)	163 (38)
Total fat, g	30 (19)	26 (17)
Food items (number of)		
Total potatoes	3 (1)	2 (1)
Total meat, pork chops	1 (1)	1 (1)
Total green peas, tbsp	2 (2)	3 (1)
Total gravy, tbsp	2 (1)	1.3 (1)
Bread, slices	1 (1)	1 (0)
Cucumber, slices of 5	1 (1)	0.83 (0.5)

[a] If not noted, reported intake is excluding intake from beverages; E%: energy percentage (in kilojoules).

Table 3. Study participants' composed meals using the two meal instruments, by body mass index (BMI) (N=51).

Composed meal	BMI < 25, median (IQR)		BMI ≥ 25, median (IQR)	
	Computer plate	Lunch plate	Computer plate	Lunch plate
Reported intake [a]				
Total energy, kJ	3320 (1361)	2726 (1101)	2881 (1080)	2734 (1017)
Total energy including drinks, kJ	3513 (1160)	3061 (1436)	3006 (1022)	2989 (1273)
Total meal, g	880 (175)	842 (235)	834 (172)	758 (243)
Total carbohydrates, E%	134 (59)	142 (42)	130 (46)	121 (67)
Total carbohydrates, g	61 (21)	58 (15)	57 (31)	47 (22)
Total proteins, E%	121 (33)	117 (21)	121 (25)	126 (29)
Total proteins, g	58 (33)	41 (32)	42 (32)	41 (32)
Total fat, E%	167 (33)	151 (29)	167 (25)	167 (38)
Total fat, g	35 (26)	26 (22)	30 (19)	26 (16)
Food items (number of)				
Total potatoes	3 (1)	3 (1)	3 (1)	2 (1.5)
Total meat, pork chops	1.5 (1)	1 (1)	1 (1)	1 (1)
Total green peas, tbsp	2 (2)	3 (1.5)	3 (2)	3 (1)
Total gravy, tbsp	1.5 (1)	1.6 (1)	2 (1)	1.3 (1.3)
Bread, slices	1 (1)	1 (0)	1 (1)	1 (1)
Cucumber, slices of 5	1 (2)	0.8 (0.7)	2 (1)	0.7 (0.7)

[a] If not noted, reported intake is excluding intake from beverages; E%: energy percentage (in kilojoules).

The quantities of the participants' chosen food items were similar between the 2 instruments. Only green peas differed, with an underestimation of 1 tablespoon when using the computer plate compared to the real lunch plate (Table 2). Using the Bland-Altman statistics, we found a tendency of agreement in mean EI within the ±10% kJ interval. More than 60% of the normal weight participants' EIs were represented in this interval. Among the overweight participants, a stronger pattern of agreement was found, with 78% of the participants' mean EIs represented in this interval (Figures 3 and 4).

Overall, our Spearman rank correlations and concordance correlations between the instruments were both equal to 0.70 for total EIs (slightly higher when including drinks in the calculations), ρ=0.59 and ρ_c=0.76 for carbohydrates, ρ=0.70 and ρ_c=0.81 for proteins, and ρ=0.58 and ρ_c=0.66 for fat (Table 4). All correlations were significant. Further, correlations for specific food items between the 2 instruments ranged from 0.46 to 0.71 for Spearman rank correlations and 0.47 to 0.72 for concordance correlations, with the lowest correlations for gravy and slices of cucumber (Table 4).

Overall, we noted somewhat higher correlations of reported number of food items for those with a BMI ≥ 25 kg/m^2 (Table 5).

Table 4. Spearman rank correlations (ρ) and concordance correlation coefficients (ρ_c) between the participants' composed meals using the two meal instruments (N=51).

Composed meal		ρ^a	$\rho_c{}^a$
Reported intake [b]			
	Total energy, kJ	0.70	0.70
	Total energy including drinks, kJ	0.79	0.72
	Total meal, g	0.72	0.68
	Total carbohydrates, kJ	0.59	0.76
	Total carbohydrates, g	0.69	0.75
	Total proteins, kJ	0.70	0.81
	Total proteins, g	0.71	0.71
	Total fat, kJ	0.58	0.66
	Total fat, g	0.63	0.65
Food items (number of)			
	Potatoes	0.65	0.72
	Meat, pork chops	0.69	0.70
	Green peas, tbsp	0.65	0.47
	Gravy, tbsp	0.48	0.50
	Bread, slices	0.71	0.67
	Cucumber, slices of 5	0.46	0.47

[a] All are statistically significant (*P*<.05).

[b] If not noted, reported intake is excluding intake from beverages.

Table 5. Spearman rank correlations (ρ) and concordance correlation coefficients (ρ_c) between participants' composed meals using the two meal instruments by body mass index (BMI).

Composed meal	BMI < 25 (n=18)		BMI ≥ 25 (n=33)	
	ρ[a]	ρ_c[a]	ρ[a]	ρ_c[a]
Reported intake [b]				
Total energy, kJ	0.76	0.63	0.67	0.73
Total energy including drinks, kJ	0.76	0.64	0.82	0.74
Total meal, g	0.68	0.70	0.70	0.66
Total carbohydrates, kJ	0.34	0.44	0.70	0.85
Total carbohydrates, g	0.49	0.60	0.77	0.78
Total proteins, kJ	0.60	0.59	0.75	0.86
Total proteins, g	0.76	0.59	0.69	0.76
Total fat, kJ	0.42	0.30	0.66	0.79
Total fat, g	0.57	0.52	0.68	0.71
Food items (number of)				
Potatoes	0.38	0.31	0.75	0.79
Meat, pork chops	0.59	0.55	0.74	0.77
Green peas, tbsp	0.70	0.41	0.65	0.53
Gravy, tbsp	0.49	0.53	0.57	0.50
Bread, slices	0.62	0.59	0.76	0.71
Cucumber, slices	0.23	0.41	0.57	0.53

[a] All are statistically significant ($P<.05$).

[b] If not noted, reported intake is excluding intake from beverages.

Higher Spearman rank correlations and concordance correlation coefficients for food items were found for those who started with the real lunch plate (0.80 and 0.61), in comparison to those who started with the computer plate, respectively. Regarding the Spearman rank correlations and concordance correlation coefficients in relation to age, we found higher coefficients for potatoes and meat for those participants who were ≥ 45 years, but higher coefficients for peas and gravy for those participants < 45 years (data not shown). Based on the questionnaire about meal satisfaction, the participants reported, on average, a level of 7 (mode 8) on the grading scale (0 = low; 10 = maximum) for fullness after meal intake.

Figure 3. Bland-Altman plot of differences between total energy intake (EI, kJ) of the computer plate and the real lunch plate against the mean of EI (kJ) for each participant with a BMI < 25 kg/m2 (n=18).

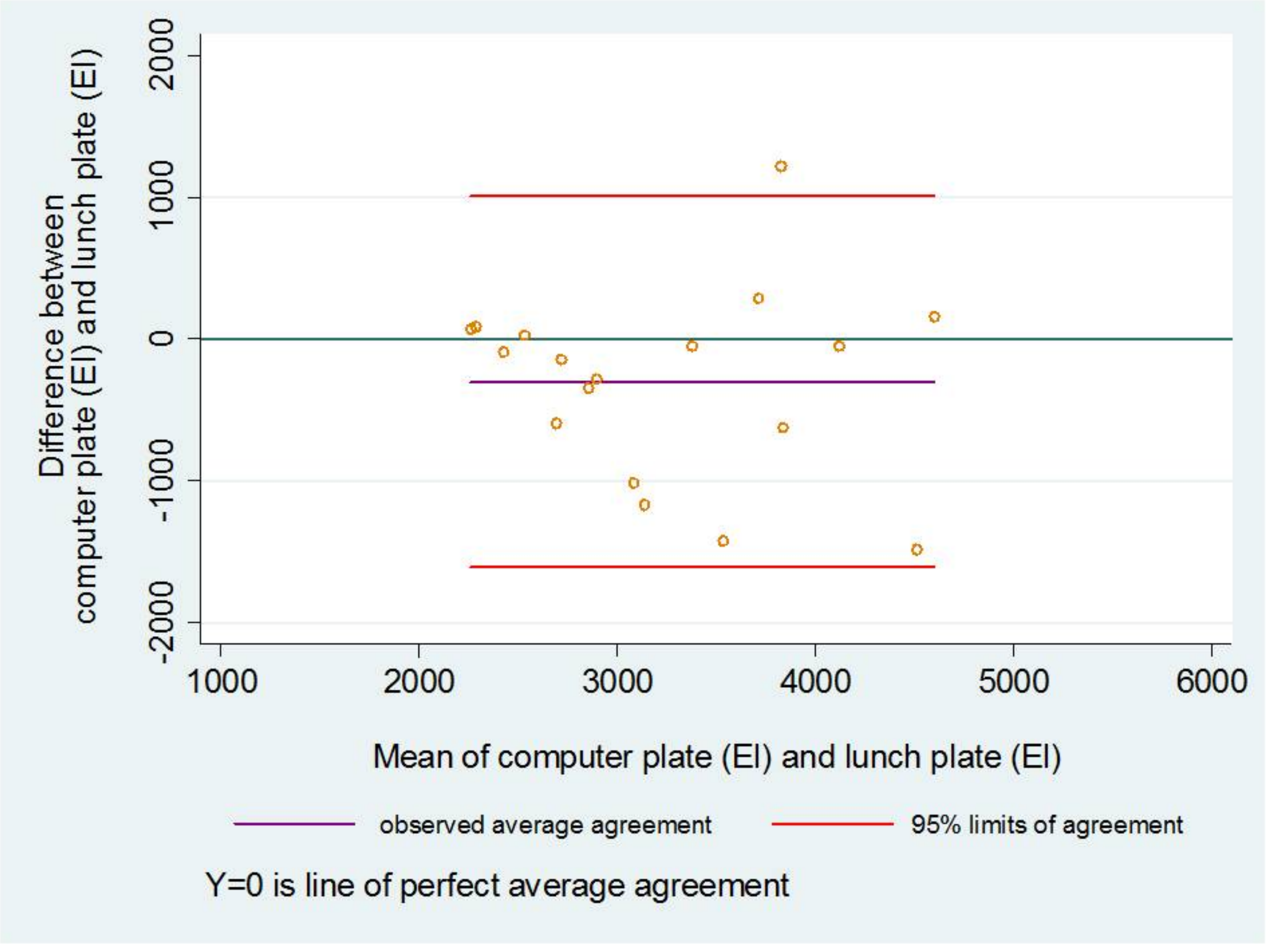
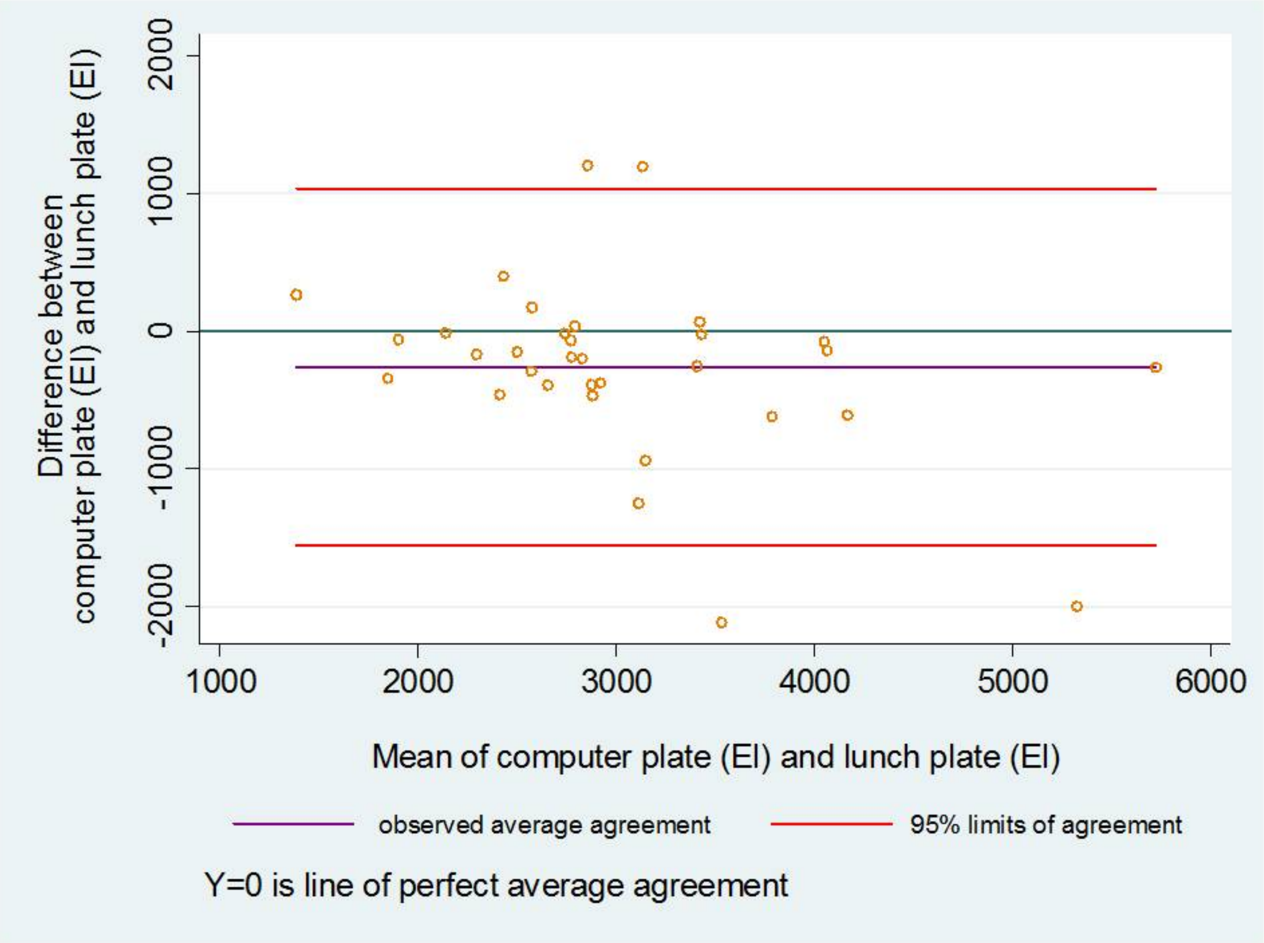


Figure 4. Bland-Altman plot of the differences between total energy intake (EI, kJ) of the computer plate and the real lunch plate against the mean of EI (kJ) for each participant with a BMI ≥25 kg/m2 (n=33).



Discussion

The results of this study support the validity of our novel interactive Internet-based food plate to measure lunch intake. The correlations between the participants' reported meal intake using the computer plate and the real lunch plate were high. Spearman rank correlations and concordance correlations, representing total reported EIs (excluding beverages and condiments), nutritional components, and food items ranged from 0.46 to 0.72 and 0.47 to 0.81, respectively. The strongest correlations were observed for protein, replicating findings from earlier studies [25].

A meta-analysis of FFQs reported Pearson coefficients ranging from 0.46 to 0.58 for EI. The authors concluded that FFQs including questions about more food items produced better estimates than FFQs with fewer items [26]. Conversely, we found strong correlations by using only a few items in a single meal. Similar correlations for specific nutrients were found in a study using cell phone cameras to document food intake [27]. However, because our instrument provides a new way of measuring food intake, the results from our study are incomparable with most previous studies.

Williamson et al [17] found higher correlations of portion sizes from direct visual estimation of a meal rather than from digital photographs. We report an overestimation using our new method on the computer where pictures of food items serve as an integral part of the meal composing process. Although reporting slightly different results, it should be noted that the participants in the study by Williamson et al estimated portion sizes based on fully prepared meals; we allowed the participants to compose a meal using suggested food items. A greater accuracy of food intake reporting has been found using more pictures of food items than fewer items when composing meals [28]. The use of several pictures of separate food items in our participants' meal composition is thus supported by previous research.

An important point of discussion is the overestimation, rather than underestimation, of food intake noted in our study. All participants overestimated their EI using the computer plate compared to the real lunch plate. Although not significant, even the overweight participants seemed to report higher EIs, contradicting previous experiences [29,30]. Overweight individuals have been found to underreport food intake, with a greater degree of underreporting with increasing BMI [31]. Yet, our results indicate higher correlations of meal intakes from our instrument for those who were overweight as opposed to our normal weight participants.

The fact that our study included only men, mostly middle-aged, and office workers are major limitations in this study preventing us from expanding our findings to participants who are women, not used to working with the computer, or those characterizing age groups other than in our study sample. Another noteworthy factor of this study is that our study sample seemed healthier (ie, much lower number of smokers) than the general population in Sweden [32], which is a common phenomenon among participants in health research [33]. A healthier lifestyle may have influenced the participants' ability to report lunch intake, and thus the overall validity and applicability of our study.

Also, we only assessed a lunch meal, including only a sample of food items available in an ordinary complete food intake. Preferably, our computer plate should measure total food intake representing various food items, meal options, and combinations. Future research is strongly recommended to explore the ability of our instrument to assess food intake in its entirety.

Even with the limitations of the current study, there are several strengths of the study design that should be highlighted. First, we validated our new measurement tool by using food items commonly consumed by the Swedish population, facilitating the participants' relatedness in the reporting process of a usual meal intake. Also, the participants were asked to rate their level of fullness after meal intake, allowing us to verify that their registered meal accounted for an actual intake. Moreover, we conducted the study in the company's dining hall and used the facility's own dishware and cutlery. In this way, the participants were familiar with the tools (eg, size of food plate, glasses), colleagues, and study climate, thus minimizing potential reporting bias and perhaps some bias from being observed by the researchers.

Although the participants performed the 2 meal assessments with a 3-week time interval to avoid recall bias [34], we noticed higher correlations of reported EI for the group who started with the real food plate. Therefore, we cannot rule out the presence of recall bias in the present study. Whether this result originates from an enhanced memory of recalling previously reported intake attributable to physically composing a lunch meal (compared to the abstract format of the virtual food plate) is difficult to state. The difference between the two assessments was that the participants ate the real lunch meal. The participants' experience from eating a meal in the lunch area may, therefore, evoke emotions and experiences contributing to an increased memory of the real lunch meal explaining the higher correlations among this group.

Overall, the results from this study demonstrate promising value in food intake assessments. The concept of our computer plate could be extended to examine estimation of daily food intakes. In addition, it may serve as a pedagogical instrument to teach healthier food habits. In fact, by taking advantage of today's advancements in technology, our computer plate could also be integrated into smartphone technology [15]. Allowing individuals to report food intake via their smartphones is a way to promote time-effective and more accurate food reporting. Future research should focus on the development of our concept, a virtual computerized food plate, to obtain a complete food intake measurement.

Conclusion

To incorporate the potential of visual and technological advancements in food intake assessments, we developed an Internet-based interactive virtual food plate to measure lunch intake. The validity of our new instrument was high, thereby producing promising applicability in health research. The concept of our computerized food plate could be further developed to assess a complete food intake.

Acknowledgments

The authors would like to acknowledge the contributions from all of the collaborating units and assistantship during the conduct of this study: our colleagues at the Obesity unit, Karolinska University hospital (a special note to Lena Svensson Mannström), and the Swedish Transport Administration's office in Sundbyberg, Sweden (particularly Nils Edström and the receptionists). The authors would also like to send a special thank you to students Judith van den Broek and Frida Alvarez for their valuable efforts during the data collection process and management of the preliminary results. Last but not least, we would like to thank Anna Westerlund for helpful feedback in the write-up of the manuscript and Stephan Rössner for advice in the planning process of the study.

The study was financially funded by the Swedish Transport Administration and the Erik and Edith Fernström's foundation, Sweden.

Conflicts of Interest

None declared.

References

1. Cade J, Thompson R, Burley V, Warm D. Development, validation and utilisation of food-frequency questionnaires - a review. Public Health Nutr 2002 Aug;5(4):567-587. [doi: 10.1079/PHN2001318] [Medline: 12186666]
2. Beechy L, Galpern J, Petrone A, Das SK. Assessment tools in obesity - psychological measures, diet, activity, and body composition. Physiol Behav 2012 Aug 20;107(1):154-171. [doi: 10.1016/j.physbeh.2012.04.013] [Medline: 22548766]
3. Poslusna K, Ruprich J, de Vries JH, Jakubikova M, van't Veer P. Misreporting of energy and micronutrient intake estimated by food records and 24 hour recalls, control and adjustment methods in practice. Br J Nutr 2009 Jul;101 Suppl 2:S73-S85. [doi: 10.1017/S0007114509990602] [Medline: 19594967]

4. Isaksson B. A critical evaluation of the duplicate-portion technique in dietary surveys. Eur J Clin Nutr 1993 Jul;47(7):457-460. [Medline: 8404781]
5. Barrett-Connor E. Nutrition epidemiology: how do we know what they ate? Am J Clin Nutr 1991 Jul;54(1 Suppl):1825-1875 [FREE Full text]
6. Bingham SA. Limitations of the various methods for collecting dietary intake data. Ann Nutr Metab 1991;35(3):117-127. [Medline: 1952811]
7. Scagliusi FB, Ferriolli E, Pfrimer K, Laureano C, Cunha CS, Gualano B, et al. Characteristics of women who frequently under report their energy intake: a doubly labelled water study. Eur J Clin Nutr 2009 Oct;63(10):1192-1199. [doi: 10.1038/ejcn.2009.54] [Medline: 19603055]
8. Goris AHC. Validation of the Assessment of Food Intake in Humans. Maastricht: University of Maastricht; 2001.
9. Heitmann BL, Lissner L. Dietary underreporting by obese individuals--is it specific or non-specific? BMJ 1995 Oct 14;311(7011):986-989 [FREE Full text] [Medline: 7580640]
10. Heitmann BL, Lissner L. [Obese individuals underestimate their food intake--which food groups are under-reported?]. Ugeskr Laeger 1996 Nov 25;158(48):6902-6906. [Medline: 8984752]
11. Høidrup S, Andreasen AH, Osler M, Pedersen AN, Jørgensen LM, Jørgensen T, et al. Assessment of habitual energy and macronutrient intake in adults: comparison of a seven day food record with a dietary history interview. Eur J Clin Nutr 2002 Feb;56(2):105-113 [FREE Full text] [doi: 10.1038/sj.ejcn.1601292] [Medline: 11857043]
12. Wansink B, Chandon P. Meal size, not body size, explains errors in estimating the calorie content of meals. Ann Intern Med 2006 Sep 5;145(5):326-332. [Medline: 16954358]
13. Illner AK, Freisling H, Boeing H, Huybrechts I, Crispim S, Slimani N. Review and evaluation of innovative technologies for measuring diet in nutritional epidemiology. Int J Epidemiol 2012 Aug;41(4):1187-1203. [doi: 10.1093/ije/dys105] [Medline: 22933652]
14. Bennett GG, Glasgow RE. The delivery of public health interventions via the Internet: actualizing their potential. Annu Rev Public Health 2009;30:273-292. [doi: 10.1146/annurev.publhealth.031308.100235] [Medline: 19296777]
15. Ngo J, Engelen A, Molag M, Roesle J, García-Segovia P, Serra-Majem L. A review of the use of information and communication technologies for dietary assessment. Br J Nutr 2009 Jul;101 Suppl 2:S102-S112. [doi: 10.1017/S0007114509990638] [Medline: 19594959]
16. Turconi G, Guarcello M, Berzolari FG, Carolei A, Bazzano R, Roggi C. An evaluation of a colour food photography atlas as a tool for quantifying food portion size in epidemiological dietary surveys. Eur J Clin Nutr 2005 Aug;59(8):923-931. [doi: 10.1038/sj.ejcn.1602162] [Medline: 15928683]
17. Williamson DA, Allen HR, Martin PD, Alfonso AJ, Gerald B, Hunt A. Comparison of digital photography to weighed and visual estimation of portion sizes. J Am Diet Assoc 2003 Sep;103(9):1139-1145. [doi: 10.1053/jada.2003.50567] [Medline: 12963941]
18. Elinder LS, Brunosson A, Bergström H, Hagströmer M, Patterson E. Validation of personal digital photography to assess dietary quality among people with intellectual disabilities. J Intellect Disabil Res 2012 Feb;56(2):221-226. [doi: 10.1111/j.1365-2788.2011.01459.x] [Medline: 21801265]
19. Black AE, Prentice AM, Coward WA. Use of food quotients to predict respiratory quotients for the doubly-labelled water method of measuring energy expenditure. Hum Nutr Clin Nutr 1986 Sep;40(5):381-391. [Medline: 3771290]
20. Hill AJ, Magson LD, Blundell JE. Hunger and palatability: tracking ratings of subjective experience before, during and after the consumption of preferred and less preferred food. Appetite 1984 Dec;5(4):361-371. [Medline: 6529262]
21. Wilcoxon F. Individual comparisons of grouped data by ranking methods. J Econ Entomol 1946 Apr;39:269. [Medline: 20983181]
22. Pagano M, Gauvreau K. Principles of Biostatistics. Belmont, CA: Duxbury Press; 1993.
23. Lin LI. A concordance correlation coefficient to evaluate reproducibility. Biometrics 1989 Mar;45(1):255-268. [Medline: 2720055]
24. Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res 1999 Jun;8(2):135-160. [Medline: 10501650]
25. Larsson CL, Westerterp KR, Johansson GK. Validity of reported energy expenditure and energy and protein intakes in Swedish adolescent vegans and omnivores. Am J Clin Nutr 2002 Feb;75(2):268-274 [FREE Full text] [Medline: 11815317]
26. Molag ML, de Vries JH, Ocké MC, Dagnelie PC, van den Brandt PA, Jansen MC, et al. Design characteristics of food frequency questionnaires in relation to their validity. Am J Epidemiol 2007 Dec 15;166(12):1468-1478 [FREE Full text] [doi: 10.1093/aje/kwm236] [Medline: 17881382]
27. Wang DH, Kogashiwa M, Kira S. Development of a new instrument for evaluating individuals' dietary intakes. J Am Diet Assoc 2006 Oct;106(10):1588-1593. [doi: 10.1016/j.jada.2006.07.004] [Medline: 17000191]
28. Subar AF, Crafts J, Zimmerman TP, Wilson M, Mittl B, Islam NG, et al. Assessment of the accuracy of portion size reports using computer-based food photographs aids in the development of an automated self-administered 24-hour recall. J Am Diet Assoc 2010 Jan;110(1):55-64. [doi: 10.1016/j.jada.2009.10.007] [Medline: 20102828]

29. Lichtman SW, Pisarska K, Berman ER, Pestone M, Dowling H, Offenbacher E, et al. Discrepancy between self-reported and actual caloric intake and exercise in obese subjects. N Engl J Med 1992 Dec 31;327(27):1893-1898. [doi: 10.1056/NEJM199212313272701] [Medline: 1454084]
30. Trabulsi J, Schoeller DA. Evaluation of dietary assessment instruments against doubly labeled water, a biomarker of habitual energy intake. Am J Physiol Endocrinol Metab 2001 Nov;281(5):E891-E899 [FREE Full text] [Medline: 11595643]
31. Johansson G, Callmer E, Gustafsson JA. Validity of repeated dietary measurements in a dietary intervention study. Eur J Clin Nutr 1992 Oct;46(10):717-728. [Medline: 1330511]
32. Swedish National Institute of Public Health. Swedish National Institute of Public Health. 2010. [Lifestyle report] URL: http://www.fhi.se/PageFiles/10796/A2010-13-Levnadsvanor-lagesrapport-2010.pdf [accessed 2012-12-31] [WebCite Cache ID 6DK3f3GO4]
33. McMichael AJ. Standardized mortality ratios and the "healthy worker effect": Scratching beneath the surface. J Occup Med 1976 Mar;18(3):165-168. [Medline: 1255276]
34. Willett WC. Nutritional epidemiology issues in chronic disease at the turn of the century. Epidemiol Rev 2000;22(1):82-86 [FREE Full text] [Medline: 10939011]

Abbreviations

BMI: body mass index
EI: energy intake
E%: energy percentage
FFQ: food frequency questionnaires
IQR: interquartile range
kJ: kilojoule
PAL: physical activity level
ρ: Spearman rank correlation
ρc: concordance correlation coefficient

Edited by G Eysenbach; submitted 18.06.12; peer-reviewed by F Slinde, S Shahar, S Pintauro; comments to author 25.08.12; revised version received 04.10.12; accepted 04.10.12; published 18.01.13.

Please cite as:
Svensson M, Bellocco R, Bakkman L, Trolle Lagerros Y
An Interactive Internet-Based Plate for Assessing Lunchtime Food Intake: A Validation Study on Male Employees
J Med Internet Res 2013;15(1):e13
URL: http://www.jmir.org/2013/1/e13/
doi:10.2196/jmir.2217
PMID:23335728

Original Paper

Evaluating the Translation Process of an Internet-Based Self-Help Intervention for Prevention of Depression: A Cost-Effectiveness Analysis

Ove K Lintvedt[1*], PsyD; Kathleen M Griffiths[2*], PhD; Martin Eisemann[1*], PhD; Knut Waterloo[1*], PhD

[1]Faculty of Health Sciences, Department of Psychology, University of Tromsø, Tromsø, Norway
[2]Centre for Mental Health Research, The Australian National University, Canberra, Australia
[*]all authors contributed equally

Corresponding Author:
Knut Waterloo, PhD
Faculty of Health Sciences
Department of Psychology
University of Tromsø
Huginbakken 32
Tromsø, N-9037
Norway
Phone: 47 77 64 52 90
Fax: 47 77 64 52 91
Email: knut.waterloo@uit.no

Abstract

Background: Depression is common and treatable with cognitive behavior therapy (CBT), for example. However, access to this therapy is limited. Internet-based interventions have been found to be effective in reducing symptoms of depression. The International Society for Research on Internet Interventions has highlighted the importance of translating effective Internet programs into multiple languages to enable worldwide dissemination.

Objective: The aim of the current study was to determine if it would be cost effective to translate an existing English-language Internet-based intervention for use in a non-English-speaking country.

Methods: This paper reports an evaluation of a trial in which a research group in Norway translated two English-language Internet-based interventions into Norwegian (MoodGYM and BluePages) that had previously been shown to reduce symptoms of depression. The translation process and estimates of the cost-effectiveness of such a translation process is described. Estimated health effect was found by using quality-adjusted life years (QALY).

Results: Conservative estimates indicate that for every 1000 persons treated, 16 QALYs are gained. The investment is returned 9 times and the cost-effectiveness ratio (CER) is 3432. The costs of the translation project totaled to approximately 27% of the estimated original English-language version development costs.

Conclusions: The economic analysis shows that the cost-effectiveness of the translation project was substantial. Hopefully, these results will encourage others to do similar analyses and report cost-effectiveness data in their research reports.

(J Med Internet Res 2013;15(1):e18) doi:10.2196/jmir.2422

KEYWORDS

Internet; Internet intervention; Cognitive Behavior Therapy; Quality-Adjusted Life Years; Cost Effectiveness; Mental health; Depression

Introduction

Computer-aided psychotherapy is a promising way to increase accessibility to evidence-based treatment of many mental disorders, such as mood disorders [1]. Depression is an important global public health issue because of its prevalence [2], that it is associated with a high level of disability and disease burden [3], and its social costs [4]. Depression is also associated with increased risk of physical disorders and early death [5,6].

In a systematic review of computer-aided psychotherapy programs, Marks et al [7] found 97 computer-aided psychotherapy programs described in 175 studies, of which 103 were randomized controlled trials (RCT). These studies discussed screening, effectiveness, efficacy, cost-effectiveness, and the dissemination of computer-aided psychotherapy programs within health services. Cohen's *d* effect sizes for the identified computer-aided psychotherapy systems ranged from 0.2 (small) to 4.3 (extremely large) [7]. For depression, Marks and colleagues [7] identified 9 programs, of which 3 were Internet-based programs evaluated in RCTs (Overcoming Depression on the Internet [8], MoodGYM [9], and netCBT [10]). In their early reviews of Internet-based RCTs of mental disorders, Griffiths and Christensen [11] and Griffiths et al [12,13] identified another program [14]. Since these reviews, at least 16 new Internet-based intervention programs for depression have been deployed. Several of these have been subjected to research evaluation or are in the test phase: Alles onder controle [15], Color your life [16], Deprexis [17], E-couch [18], HealthSteps for Depression [19], Interapy Depression [20], Living Life to The Full On-line [21], MoodCalmer [22], MoodHelper.org [23], MoodMemos [24], myCompass [25], Project Catch-it [26], This Way Up-Depression Course [27], This Way Up-Mixed Depression and Anxiety Course [28], Xanthis [29], and Youth Mental Health [30]. In addition, more than 20 other programs are still being tested and, thus, no published research evidence of the efficacy of the programs are available.

A recent quantitative meta-analysis on Internet and computerized interventions for adult depression found an overall effect size of $d=0.41$ for computer-aided psychotherapy compared to control [31]. A recent meta-analysis by Andrews and colleagues [32] found an overall effect size for depression of Hedges' $g=0.78$. Most of the research on Internet-based interventions for depression has focused on treatment rather than prevention, and most studies have employed guided rather than pure self-help strategies. One reason for this focus could be that treatment has more immediate benefits than prevention in which the benefits take longer to emerge [33].

In their meta-analysis, Andersson and Cuijpers [31] reported an average effect size of $d=0.25$ for unguided Internet and computerized interventions. It has been suggested that such interventions have higher attrition rates than those involving therapist support [31,34-36]. To date, we are not aware of any published studies that directly compare attrition in therapist-guided compared to unguided interventions. However, even if attrition is higher in unguided Internet interventions, they have the potential to provide significant public health benefits due to the degree of dissemination [33].

Internet-based self-help has the potential to reach target groups with an unmet need for help. In a recent study, we found that two-thirds of the participants completing a trial of an Internet-unguided cognitive behavioral therapy (CBT) program initially reported an unmet need for help with a psychological problem [37]. Briefly, the Internet intervention was associated with a significant reduction in depressive symptoms compared to the control condition, and thus indicated that the intervention facilitated an effective self-help effort among a group of people who normally would not seek professional help. The importance of reaching and engaging this group is obvious given the evidence that low levels of help-seeking intentions among those with mental health problems are associated with suicidal ideation [38].

There is now sufficient evidence to suggest that Internet interventions can be effective in the prevention and treatment of depression [39,18,40,41]. The International Society for Research on Internet Interventions (ISRII) has highlighted the importance of facilitating the dissemination of Internet applications by providing translation into multiple languages [42]. The World Health Organization (WHO) has also suggested that after establishing their efficacy, Internet-based prevention interventions should be disseminated worldwide [33]. One program that has been subjected to a number of trials and found to be effective over 12 months is the English-language Internet program for depression, MoodGYM [43]. There is also evidence of the efficacy of the English-language website BluePages [44], which provides information on the symptoms of and treatment for depression based on scientific evidence, help resources, and depression and anxiety screening tests. Both programs are available as part of a publicly available, free of charge, e-mental health service delivered by The Australian National University (ANU) and funded by the Australian Government [45]. This paper describes the translation of MoodGYM and BluePages into the Norwegian language by a research group at the University of Tromsø in Norway. Reviews in the field of Internet interventions for mental disorders show that studies are lacking cost-effectiveness data [11,46].

The aim of this study is to evaluate the translation process of an Internet-based self-help intervention up to dissemination and to study cost and cost-effectiveness of the project. It will also consider the feasibility of providing access to Internet-based interventions in the national language versions. Further, we discuss some of the challenges experienced during this project and give recommendations for future developments. Finally, we attempt to demonstrate the use of disease-specific scales to estimate quality of life.

Method

Prior Work

Intention to Use Internet-Based Self-help

Before translating the Internet-based interventions, a study was undertaken to investigate the need for a Web-based self-help intervention [47]. Nearly 31.9% (115/367) of the respondents reported they felt a need for help, but had not sought help; therefore, they had an unmet need for help. Among these respondents, 91.2% (105/115) reported a positive attitude toward using a service like MoodGYM.

The Norwegian Version of MoodGYM and BluePages

In February 2006, Norwegian language versions of BluePages and MoodGYM were made available. The planning of this project started in July 2004 and the project was completed in June 2006. The translation of BluePages was formalized in a licensing deed, specifying ANU as the intellectual property owner and the obligations for the collaborating partners. For MoodGYM, the research collaboration agreement specified a nonexclusive, nontransferable license for the translated version during the period of collaboration and joint ownership of the data emerging from the collaboration. The translation of MoodGYM and BluePages was carried out between October 2005 and January 2006. The Norwegian version of the BluePages website was developed by using a 2-phase process. First, a professional translator prepared a Norwegian version of BluePages. In the second phase, the translation was adjusted by the research group at the University of Tromsø in Norway to ensure that it was culturally and clinically appropriate for Norwegian users. The MoodGYM training program was translated in 4 phases. The research group conducted the first phase of translation. In phase 2, clinical professionals with formal competence in cognitive therapy scrutinized the translation from the first phase and made adjustments to the text when necessary. In phase 3, an expert translator of English compared the Norwegian version of the program with the original Australian version and checked for inconsistencies. In the fourth phase, the research group evaluated all changes and finalized the translation.

Validating the Norwegian Version

The Norwegian versions of MoodGYM and BluePages were evaluated in a RCT that compared the effect on depressive symptoms of an unguided Internet-based intervention. The Internet condition consisted of a depression information website and a self-help Web application that delivered automated CBT. The participants in the waiting list condition were free to access formal or informal help as usual. This trial was organized as an unguided quasi-indicated prevention intervention. A total of 163 students (mean age 28.2 years) with elevated psychological distress were enrolled into the trial. The intent-to-treat effect size for depressive symptoms was d=0.63 and d=0.72 for completers [37]. A total of 61.9% (101/163) of the participants remained in the trial at postintervention, including 53% (43/81) of the Internet intervention participants. With respect to adherence, 33% (27/81) of the Internet intervention participants completed on average 63% (3.1/5 modules) of the MoodGYM program without any support or reminders. Among the completers, 63% (27/43) in the experimental condition and 56% (33/59) in the control condition initially reported an unmet need for help. The logs from MoodGYM showed that the 25% of the 81 participants in the experimental condition who did not return the postintervention questionnaire, on average had used almost half of the modules in MoodGYM (2.4/5 modules, 48%). The trial demonstrated a real-world usage of an unguided Internet-based intervention.

Cost Estimates

The total cost estimate for the Norwegian version of MoodGYM and BluePages was estimated from the cost of the translation process and the project costs for the Norwegian evaluation trial.

The translation process was accomplished by using 2 members from the Norwegian research group, 3 students in their final year of clinical psychology training, and external resources (translators, psychiatrists, mailing/printing services). Translation costs for employees and students were calculated based on the estimated time consumed and their salaries. The cost for external resources was based on payments for their services.

The trial costs were incurred over 8 weeks from initial participant contact until the completion of the validation trial. This estimate for total costs was based on all expenditures of the validation trial for BluePages and MoodGYM and included the direct costs of mailing and printing during the trial, and also salary for employees and students. Working hours for developing funding applications were not included in the estimates. Because the aim of this study was to evaluate the translation process up to dissemination, the trial costs are included even though these costs could be regarded as research costs.

The initial cost data were in Norwegian kroner (NOK) and dated back to late 2005 and early 2006. For the purpose of this paper, the cost has been adjusted for inflation based on the Norwegian consumer price index into a 2009 price level [48] at an average annual inflation rate of 2.3%. The development cost from ANU dated back to 2001. We have adjusted the cost for inflation based on the Australian consumer price index into a 2009 price level [49], at an average annual inflation rate of 2.9%. The foreign currencies were then converted into Euro (€) based on real exchange rates from January 1, 2010. These rates form the basis of our calculations, rounded up to the nearest €100. The rationale for using the Australian development cost was in the interest of estimating the development cost if the project had been developed from scratch in Norway, and the Australian development cost provided an estimate of this figure. The cost analyses were based on an operating period of 3 years. After such a timeframe, the Internet-based programs need to be maintained, changes made in design, the technology needs to be updated, and so on. Although most of the investment in the intervention can be maintained, there will be some new costs at this point in time.

Measures

Center for Epidemiologic Studies Depression Scale

The Center for Epidemiologic Studies Depression Scale (CES-D) [50] was used as the outcome measure for depressive symptoms. Each item of the CES-D is scored on a 4-point ordinal scale ranging from 0 to 60, with a total score of 16 or higher considered to be depression. Because this criterion has some limitations [51], it will not be used in this paper. For the completers analyses, we used a stratified scale adapted from Rushton and colleagues [52]: subclinical (0-15), mild/moderate (16-23), and moderate/severe (≥24) depression. The internal consistency of the CES-D scale was alpha = .87.

Kessler Psychological Distress Scale

The Kessler Psychological Distress Scale (K10) [53] is a scale of psychological distress developed for use in epidemiological surveys. Scores on the K10 scale range from 10 to 50, with higher scores indicating greater distress. People who score under 20 are likely to be well, whereas those scoring between 20 and 24 have a mild mental disorder, those between 25 and 29 have a moderate mental disorder, and those scoring 30 and over have a severe mental disorder. The internal consistency of the K10 scale was alpha = .79.

Quality of Life

There has been a growing interest for applying measures for health-related quality of life (QOL) to evaluate health care services, in general. However, in the field of Internet- and computer-based self-help for mental disorders, reviews find that studies lack cost-effectiveness data [11,46]. This is a disadvantage for the field because proper evaluations among interventions and services are not possible. There are several measures available for this purpose: those specific to groups of diseases, and generic measures applied across disease categories [54]. Some produce a single measure as a health index and others assess a profile of scores in different areas of disease. For independent interventions in which it is essential to use average cost-effectiveness ratios and for mutually exclusive interventions, the use of incremental cost-effectiveness ratios is essential [55]. Interventions in which the costs and effects are not affected by the introduction of other interventions are independent interventions. The cost-effectiveness ratio measures how efficiently the intervention can produce an additional quality-adjusted life year (QALY). In this way, the cost-effectiveness of alternative innovations may be compared. The cost-effectiveness ratio (CER) is calculated as the costs of intervention divided by the health gain (eg, life-years gained). The intervention with the lowest CER is preferred [55].

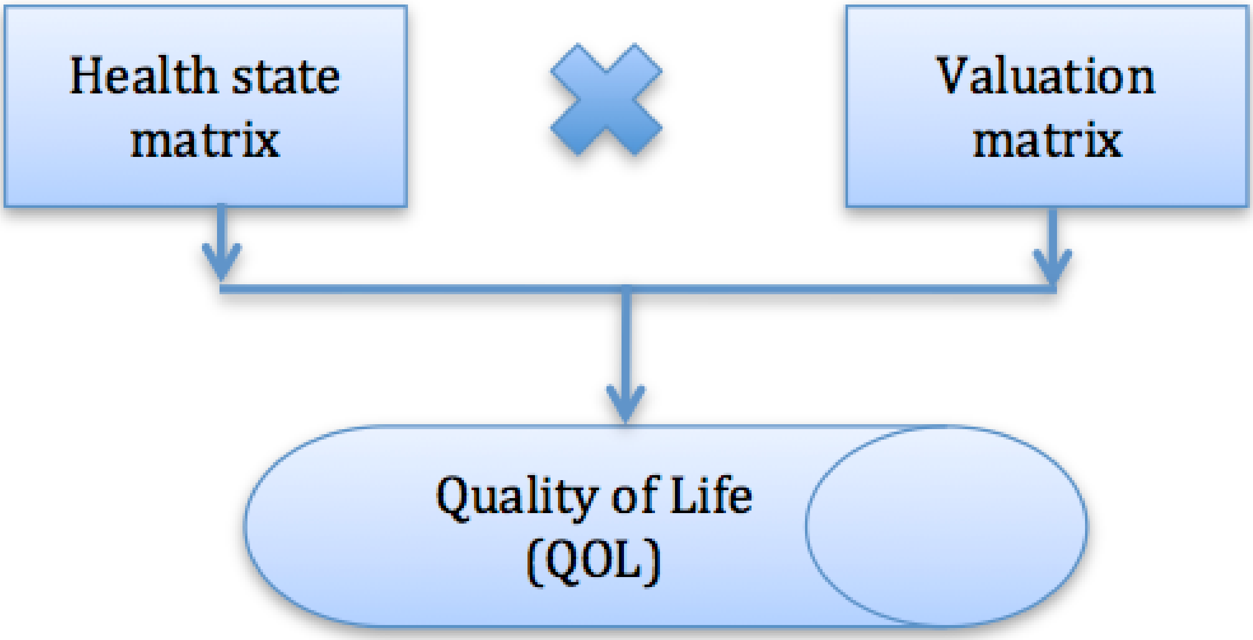
The Rosser Classification of Illness States Scale

The Rosser classification of illness states scale (Rosser Index) [56] offers a ratio scale based on QOL for estimating a utility value that captures the degree of improvement in health (ΔH). The ΔH is used to calculate the gain in QALY. As one of the earliest utility instruments, the Rosser Index was developed for use in clinical settings [57]. This is a generic measure with a single index measure for health status. The value for each health state is the sum of all products when multiplying each cell in the Rosser health state matrix with corresponding cells in the Rosser valuation matrix (Figure 1).

In the present study, the Rosser Index was used as a generic instrument to evaluate the observed effects for the Internet intervention. The method is based on classifying outcome data from the trial into 2 components: disability categories (I-VIII) and distress categories (A-D) that define 29 potential health states [56]. The Rosser health state matrix is based on these disability and distress categories.

The original valuation matrix was based on 70 respondents (doctors, nurses, patients, etc) and was not a random sample of the population. The matrix has been transformed and validated in several publications [58]. The Rosser valuation matrix in its present form is based on the method of magnitude estimation [59] in studies from England [56], but they correspond well with Norwegian data [60]. The process of deciding the appropriate categories for the Rosser health state matrix is according to the method described by Gudex and Kind [61].

Figure 1. Model for the Rosser Index measuring quality of life (QOL). The health state matrix is based on the Center for Epidemiologic Studies Depression scale (CES-D) scores in relation to the Rosser Disability Category and the Kessler Psychological Distress Scale (K10) scores in relation to the Rosser Distress Category.



Rosser Health State Matrix: Disability Categories

The CES-D was used as the basis for deciding the appropriate disability categories. No other studies were found that used CES-D as the disability categories in the Rosser health state matrix. The authors have functioned as a reference group of clinical psychologists and found that the CES-D seems to correspond to the Rosser disability dimensions (see Table 1). This is in-line with the recommendations by Gudex and Kind [61]. The first 2 CES-D subgroups (subthreshold and mild/moderate) correspond well with Rosser´s disability categories I (no social disability) and II (slight social disability); the third CES-D subgroup (moderate/severe) corresponds well when dividing the CES-D scores over Rosser´s disability categories III to VI. The disability categories VII to VIII are not used because they represent states of being confined to bed and unconscious. The analogy between CES-D and Rosser Disability category is shown in Table 1.

Table 1. Center for Epidemiologic Studies Depression scale (CES-D) scores and their corresponding Rosser disability categories.

CES-D		Rosser disability category	
Subgroup	Score	Category	Description
Subclinical	0-15	I	No social disability
Mild/moderate	16-23	II	Slight social disability, can continue almost as usual with occupational and home activities
Moderate/severe	24-33	III	Severe social disability and/or slight impairment of performance at work
	34-42	IV	Choice of work or performance at work very severely limited
	43-52	V	Unable to undertake any work/education
	53-60	VI	Full-time care or in an institution

Rosser Health State Matrix: Distress Categories

The K10 was developed for use in epidemiological surveys and is used to measure psychological distress. No other studies were found that used K10 as the distress category in the Rosser health state matrix. The authors have functioned as a reference group of clinical psychologists and found that the K10 seems to correspond largely with the Rosser distress dimensions (see Table 2). This is in-line with the recommendations by Gudex and Kind [61]. The K10 preassessment scores are used as the basis for deciding the appropriate distress categories for the participants.

Table 2. Kessler Psychological Distress Scale (K10) scores and their corresponding Rosser distress categories.

K10		Rosser distress category	
Subgroup	Score	Category	Description
Being well	10-19	A	No distress
Mild mental disorder	20-24	B	Mild distress
Moderate mental disorder	25-29	C	Moderate distress
Severe mental disorder	30-50	D	Severe distress

Quality-Adjusted Life Years

A QALY is a year of life adjusted for its quality or its value. A year in perfect health is considered equal to 1 QALY [62]. The value of a year in ill health would be discounted. In the trial, there was a period of 8 weeks between the preassessments and postassessments, in such a way that the ΔH was gained over a 2-month timeframe. The trial did not have a 12-month follow-up, but research from the field shows that the effect over time, from posttest to 12-month follow-up, is maintained [63]. Mackinnon and colleagues [63] conducted a study on 12-month outcome for MoodGYM compared with BluePages and a control group. This gives the best estimate for how to transform our findings to a 12-month outcome. They reported that both the MoodGYM group and the control group had a decline in CES-D scores from posttest to 12-month follow-up, and the control group had the best benefit in reduced symptoms of depression. These transformations of QOL are according to the principles described by Gudex and Kind [61]. The gain in QALYs (QALY-gain) is based on the utility value for improved health ΔH, multiplied by the time interval (T) over which the improvement occurred (measured in years): QALY-gain = $\Delta H \times T$ [64].

Cost per QALY Gained

The National Institute for Health and Clinical Excellence [65,66] in the United Kingdom considers cost per QALY gained among its criteria for coverage recommendations to the National Health Service. This has been converted into an explicit criterion of £30,000 per QALY to be used as a guideline for these recommendations [67]. In Norway, an amount of NOK 500,000 has been proposed (equivalent to €67,100 in 2009) as a temporary best estimate of a QALY [68].

Results

The Norwegian real cost totalled to €39,900, and the estimated cost to €56,000, adding up to a total of €95,900. Based on data published by Butler and colleagues [69], the Australian development cost was €479,400 plus an average annual maintenance cost of €44,700. The Norwegian costs correspond to almost 27% of the Australian development costs.

Table 3 summarizes the complete cases analysis for depressive symptoms with change scores and effect size [37]. The between-group difference for the Internet intervention and control groups for the overall complete cases analysis was 7.1 points on the CES-D scale and 7.0 points for the intention-to-treat analysis. The mean scores at baseline for K10 were 33.2 (SD 5.5) for the Internet group and 33.2 (SD 5.0) for the control group. The deterioration in the control group is evident in their total scores and rather high numbers in the CES-D subclinical and moderate/severe subgroups. There is almost no change for the mild/moderate subgroup.

Table 3. Completers between-group effect size (Hedges' *g*) for depression level as measured by the Center for Epidemiologic Studies Depression Scale (CES-D).

CES-D scale and condition		Pretest	Posttest	Contrast[a]	Effect size[b]
		Mean (SD), n	Mean (SD), n	Mean (SD)	*g* (SE)
Total					
	Intervention	22.6 (10.9), 43	18.5 (14.0), 43	4.1 (10.4)	0.72 (0.21)
	Control	18.5 (9.6), 59	21.4 (13.0), 59	–3.0 (9.1)	
Subclinical					
	Intervention	11.4 (3.6), 14	10.0 (8.6), 22	1.4 (8.4)	0.68 (0.34)
	Control	9.8 (3.4), 25	14.4 (9.7), 25	–4.6 (8.8)	
Mild/moderate					
	Intervention	19.0 (2.7), 10	11.0 (9.5), 7	8.0 (9.8)	0.85 (0.42)
	Control	19.2 (2.3), 17	19.5 (10.9), 11	–0.3 (9.2)	
Moderate/severe					
	Intervention	32.7 (7.0), 19	28.6 (12.8), 14	4.1 (11.7)	0.67 (0.34)
	Control	30.4 (10.0), 17	32.4 (12.1), 23	–3.2 (9.2)	

[a] Positive contrast represents better outcome.

[b] Hedges' *g* is Hedges' unbiased effect size; SE is Hedges' unbiased standard error.

The following analysis is based on the completers data from the effect trial [37]. For both conditions (Internet intervention and control group), participants' distress and disability category is given. Table 4 presents an overview for condition and time in which each cell in the matrix represents the number of cases for the combination of distress and disability. Each cell in this descriptive matrix is associated with a corresponding weight in the Rosser valuation matrix (Table 5).

Table 4. Descriptive matrix with number of participants for each disability and distress state, for condition and test time.

Condition	Disability category	Distress category, n							
		Baseline				Posttest			
		A	B	C	D	A	B	C	D
Internet	I	1	8	5		1	12	8	1
	II		6	3	1		4	2	1
	III		3	4	3		2	1	2
	IV		2		5		1	1	5
	V				2				2
Control	I	1	11	10	3		10	12	3
	II		7	4	6		6	3	2
	III		1	7	6	1	3	1	5
	IV		1				1	5	3
	V		1		1				3
	VI						1		

Table 5. Rosser´s utility valuation scores for health conditions.

Disability category	Distress category[a]			
	A	B	C	D
I (no social disability)	1.000	0.995	0.990	0.967
II (slight social disability)	0.990	0.986	0.973	0.932
III (severe social disability)	0.980	0.972	0.956	0.912
IV (severely limited work performance)	0.964	0.956	0.942	0.870
V (unable to work/study)	0.946	0.935	0.900	0.700
VI (total social disability)	0.875	0.845	0.680	0.000

[a] 1 = healthy; 0 =dead. Table shows only relevant disability categories (for complete table see Kind et al [59]).

When all cells in the descriptive matrix are multiplied with the corresponding weight in the valuation matrix, the sum of all these products gives the total QOL score for condition and time. The total QOL score is then divided by the number of participants in the group, which gives an average QOL gain (ΔH) score for the group [61]. The Internet intervention group (n=43) had a total QOL at pretest of 40.82 (mean 0.95), and 41.02 (mean 0.95) at posttest, with a ΔH of 0.005. For the control group (n=59), the total QOL at pretest was 56.86 (mean 0.96) and 55.96 (mean 0.95) at posttest, with an average ΔH of -0.015. The between-groups gain in QOL (ΔH) was 0.020.

Prediction of a 12-Month Outcome

By using the change scores from the Australian 12-month follow-up study of MoodGYM and BluePages [63], we calculated a factor to extrapolate our findings to predict a 12-month outcome. Because the results for the control group in our trial were negative, the factors were calculated from posttest to 12-month follow-up. For the MoodGYM group, the change from posttest (mean 15.9) to 12-month follow-up (mean 14.1) was 1.8 points on the CES-D scale, giving a change factor of 1.127. For the control group, the change from posttest (mean 19.5) to 12-month follow-up (mean 16.4) was 3.1 points on the CES-D scale, giving a change factor of 1.19. As the change for the control group was negative, we multiply with the inverse factor at 0.84 (1/1.19). In this way, the difference between groups will diminish.

The gain in QOL (ΔH) for the Internet intervention group was 0.0048, and increased to 0.0054 (0.0048×1.127) after transformation. The negative gain in ΔH for the control group was –0.015 and decreased to –0.013 after transformation. The between-groups ΔH reduced to 0.018. The transformation extrapolates the results to a timeframe of 1 year. This gives a QALY-gain=0.018×1, which equals 0.018 of a QALY.

Table 6 presents cost analyses assuming an operating period of 3 years, as the Internet-based programs need to be maintained after some time. The total cost for the translation project is €55,000 per year, equal to 27% of the cost of the development project in Australia. For the translation project, we added an annual cost for maintenance (€14,000) and a service fee (€9000) to provide a more realistic estimate for real-world use. The Internet-based programs run on a shared server in Australia, resulting in the lower maintenance cost for the translation project.

Table 6. Estimated development and translation costs with mean cost (based on 3 years’ operating time).

Project and costs type	Costs per year (€1000)			Mean cost (€1000)
	Year 1	Year 2	Year 3	
Development cost (Australia)				
Development	479.4			159.8
Maintenance	44.7	44.7	44.7	44.7
Total	524.1	44.7	44.7	204.5
Translation cost (Norway)				
Development	95.9			32.0
Maintenance	14.0	14.0	14.0	14.0
Service fee	9.0	9.0	9.0	9.0
Total	118.9	23.0	23.0	55.0

Based on data from the validation study, the QALY-gain was found to be 0.018 of a QALY. To obtain one QALY, we need 56 individuals (1/0.018) to use the Internet intervention in the same way as the group of completers did in the trial. This is equal to gaining 1 year of full health for 1 person.

Table 7 shows a sensitivity analysis for cost and savings, divided on a development and translation project. Estimated savings per QALY is varied between the Norwegian estimate and the UK estimate. For each of these scenarios, Table 7 shows potential savings based on the number of people treated annually. The CER is the ratio between cost and QALY. The development cost for the translation project is less than the cost for 1 QALY gained, and for the developments project just over 3 QALY gained, based on the Norwegian estimate for cost per QALY gained. The most conservative estimate indicates that for every 1000 persons treated, 16 QALYs are gained. The investment is returned 9 times and the CER is 3432 (based on the more conservative UK estimate for cost per QALY gained). Consequently, if 20,000 people are treated per year, 355 QALYs are gained. Annual savings for the translation project are estimated to at least €500,000 per 1000 persons treated.

Table 7. Estimates for cost and annual savings (based on 3 years' operating time) for development and translation projects, with cost-effectiveness ratio (CER).

Project estimate	Per QALY		Annual					
	Savings (€1000)	Persons	Development cost (€1000)	Persons treated	Savings per person[a] (€1000)	Total savings (€1000)	QALYs[b]	CER[c]
Development								
Norwegian	67	56	204	1000	0.99	992	15	13,772
Norwegian	67	56	204	20,000	1.19	23,725	354	576
UK	30	56	204	1000	0.33	332	11	18,450
UK	30	56	204	20,000	0.53	10,510	350	582
Translation								
Norwegian	67	56	55	1000	1.20	1141	17	3228
Norwegian	67	56	55	20,000	1.20	23.875	356	154
UK	30	56	55	1000	0.54	481	16	3432
UK	30	56	55	20,000	0.54	10,659	355	155

[a] Savings per person = (savings per QALY)/(persons per QALY)–(development cost)/(persons treated).

[b] QALYs = (total savings)/(savings per QALY).

[c] CER = (development cost)/QALYs.

Sensitivity Analysis

Our model for estimating the gain in QOL is based on the Rosser Index. Depending on how well the CES-D and K10 scales were fitted to the disability and distress categories, this could affect the output from the model. A sensitivity analysis was conducted to explore how the operationalizing of the scales could have contributed to variation in the outcome (QOL gain). Three scenarios were investigated: (1) the main analysis based on the score for each individual participant, (2) by using the CES-D subgroup mean, and (3) by using the mean for the conditions. In addition, 2 scenarios were investigated to explore the effect of the deterioration in the control group: (4) if there was no change in the control group, and (5) a positive change in the control group. Scenario 1 is already described previously.

Scenario 2 had only 3 values for each condition and time. In the Internet condition, scores for CES-D and K10 placed the CES-D subgroups as follows at pretest and posttest: subclinical group (disability I/distress B, K10 mean 22.8; disability I/distress B, K10 mean 23.7), the mild/moderate subgroup (disability II/distress B, K10 mean 24.4; disability II/distress C, K10 mean 25.3), and the moderate/severe subgroup (disability III/distress D, K10 mean 30.0; disability III/distress D, K10 mean 30.7). In the control condition, scores for CES-D and K10 placed the CES-D subgroups as follows at pretest and posttest: subclinical group (disability I/distress B, K10 mean 24.5; disability I/distress C, K10 mean 25.4), the mild/moderate subgroup (disability II/distress C, K10 mean 27.5; disability II/distress C, K10 mean 25.1), and the moderate/severe subgroup (disability III/distress C, K10 mean 28.5; disability III/distress C, K10 mean 28.3).

Scenario 3 was placed in the disability category II for both conditions, based on CES-D scores (Table 3). The K10 mean was 33.2 for both conditions. The same participants were present at both pretest and posttest, making the distress category D at both the pretest and posttest. When looking at the changes in CES-D scores, it is reasonable to assume that the Internet condition, with a reduction of 4.1 points, resulted in a relief of distress. The opposite could be applied for the control condition, as they had an increase of 3.0 points. The manual for the Rosser Index offers no rules for this scenario, so it seems reasonable to apply a reduction corresponding to half of the change between distress categories C and D, at 0.02. The same applies for the control condition.

Scenario 4 aimed to explore the effect of the deterioration in the control group. If there were no change in the control group during the trial, the QOL gain for the intervention would be equal to the gain for the Internet intervention group. The QOL gain between groups is 0.006, only one-third of the QOL gain from scenario 1.

Scenario 5 represents a scenario in which the control group had a positive change from baseline to posttest, as one could expect during a time period of 12 months. This scenario uses the pretest assessment for both conditions to calculate a new factor to extrapolate our findings to predict a 12-month outcome. The extrapolating was done directly on the pretest scores for each participant to predict outcome scores at 12 months. The rationale for exploring this is based on the fact that the Norwegian and Australian trials were different, unattended vs minimal contact/attention placebo. The change factors for the Internet intervention and control group were 1.546 and 1.317, respectively. The QOL gain between groups is 0.012, ie, smaller than for scenario 1 (two-thirds), but still notable. An alternative approach for this scenario could be to extrapolate the Norwegian posttest scores and compare to the Australian improvement for the control group. The change factors for the Internet intervention and control group would then be 1.127 (as for scenarios 1-3) with a QOL gain of 0.003. This is only one-sixth of the QOL gain from scenario 1. See Table 8 for the results from the sensitivity analysis.

Table 8. Quality of life (QOL) sensitivity analysis within and between conditions.

Scenario	Condition	ΔH[a]	ΔH_{ext}[b]
1. Main analysis			
	Internet QOL gain	0.005	0.006
	Control QOL gain	–0.015	–0.013
	Between-conditions QOL gain	0.020	0.018
2. Subgroup mean			
	Internet QOL gain	0.008	0.010
	Control QOL gain	–0.004	–0.003
	Between-conditions QOL gain	0.012	0.013
3. Condition mean			
	Internet QOL gain	0.020	0.024
	Control QOL gain	–0.020	–0.017
	Between-conditions QOL gain	0.040	0.041
4. Main analysis, unchanged control group			
	Internet QOL gain	0.005	0.006
	Control QOL gain	0.000	0.000
	Between-conditions QOL gain	0.005	0.006
5. Main analysis, gain in control group			
	Internet QOL gain		0.024
	Control QOL gain		0.012
	Between-conditions QOL gain		0.012

[a] ΔH=QOL gain.

[b] ΔH_{ext} = QOL gain extrapolated.

Discussion

Conservative estimates indicate that for every 1000 persons treated, 16 QALYs are gained. The investment is returned 9 times and the CER is 3432. The costs of the translation project totalled to approximately 27% of the estimated original English-language version development costs.

Principal Results

The scenarios in the sensitivity analyses differ in that they decrease in distribution among the 29 potential health states. Scenario 1 uses up to 15 different states, whereas scenario 2 uses up to 5 different states, and scenario 3 only 1 health state. As the distribution becomes more even, the QOL gain (ΔH) between groups decreases, and for scenario 3 there is no gain at all. This is partly an effect of the negative results for the control condition. If the control group is kept stable (no change between pretest and posttest score), the QOL change between groups is equal to the change within the intervention condition. Then there is an opposite effect, in which the QOL gain increases with the decrease in distribution. The result is less sensitivity and gives an inflated result. Between scenarios 1 and 2, this is probably a result of less variance and not necessarily that we had to use judgment, as in scenario 3. The results for the extrapolated QOL gain (ΔH_{ext}) offer no additional insight because it is just a fixed factor that scales the scores into a 12-month QOL gain.

Scenarios 4 and 5 explore the effect of the deterioration in the control group, as one could expect some improvement in that group over a longer time period. The Australian trial is not comparable to the Norwegian because the experimental and control groups are given different focus (minimal contact vs unattended). As scenario 4 shows, we need 3 times as many persons treated to gain 1 QALY. In scenario 5, we need 1.5 as many persons treated to gain 1 QALY.

As the economic analysis shows, the cost of the translation project is less than 1 gained QALY. In the worst case, it is necessary to treat 46 subjects with symptoms of depression to reach the cost break-even point. The treatment received should, on average, be the same as the average for the completers in the trial. As a translation project, this investment is highly cost-effective as it returns the investment many times. With the Norwegian estimate for gained QALY (savings per QALY), every 1000 persons treated yields 17 QALYs and returns the investment 21 times with a CER at 3228. Based on the UK estimate for gained QALY, every 1000 persons treated gives 16 QALYs, returns the investment 9 times, and the CER is 3432. The cost for the translation projects is less than 1 QALY, so the annual number needed to treat to make the cost break-even is 46 individuals, and the development project needs 171 individuals (based on Norwegian estimate for savings per QALY). This is promising, as we know that people, in general, are positive to Internet-based self-help and that this intervention can serve as a prevention intervention. Results from a meta-analysis show that preventive interventions can reduce the incidence of depressive disorders by 22% [70]. The authors concluded, for people with subthreshold symptoms or with high-risk situations, preventive interventions might be exactly what they need.

The trial costs are included in development costs, and include testing the Web-based interventions' effect and effectiveness. The WHO suggests that Internet-based prevention interventions should be disseminated after establishing their efficacy [33]. We consider the process of "establishing efficacy" as a part of the project and recommend others to include trial costs as well.

We regard the assumptions underlying the cost-effectiveness scenarios as conservative, but at the same time it depends on the kind of marketing, that the potential users find the interventions acceptable, and can trust the service provider. Internet-based interventions are costly to develop and require long-term research to test and evaluate their efficacy and effectiveness before they can be disseminated en masse. Locating funding sources for dissemination is often difficult [33]. This paper shows the potential economic benefits of translating Internet-based interventions into other languages, and that the cultural and linguistic transformation of tools developed in other countries are feasible. Hopefully, this paper also contributes to more research reporting cost-effectiveness data in the future. Our findings should encourage others to undertake similar translation activities. These could comprise other mental disorders for which evidence-based interventions exist, including social anxiety disorder, panic disorder, posttraumatic stress disorder, substance misuse, and eating disorders [71-75]. For small countries, like Norway, with a limited number of researchers in the e-mental health field, it is considerably more cost-effective to translate and implement existing interventions than to develop new ones from scratch. This should enable the relatively rapid introduction and dissemination of new interventions to a broad spectrum of the target population. In addition, such an approach could facilitate global dissemination to reduce health disparities across countries and cultures [76,33]. Although we find it considerably more cost-effective to translate existing interventions than to develop new ones from scratch, there could be situations in which the latter option becomes necessary or appears as more beneficial, ie, if for our area of interest no proven effective Internet-based intervention is available, the content of the program does not use the treatment we are interested in, we want to have control over the content and development of the intervention, and if there is no willingness from intervention developers to give access to their programs.

Internet-based treatments are more accessible than therapist-delivered treatments and, therefore, have the potential to reach more people with an unmet need [77]. An important further step is to assess the cost-effectiveness of the various Internet-based treatments to allow policy makers and health care providers to compare the benefits of online treatment with other interventions. This paper represents one of the few demonstrating the cost-efficacy of an Internet-based therapeutic intervention to treat or prevent a mental disorder and the only study of the cost-effectiveness of translating an existing program for use in another language, as far as we know.

Although the current study demonstrates that the translation of interventions can be cost-effective, there are other issues that need to be addressed in establishing a translation agreement. Firstly, the licensing deed or contract will need to specify the degree to which the translator can adapt and customize the translated version to local requirements. Secondly, this agreement will need to specify the extent to which the translating site will have the opportunity to implement new functionalities, such as customized printouts. Thirdly, the agreement will need to specify the extent to which the translating site will have an opportunity to influence and contribute to new content for the program. If the content of the application is altered, the application is no longer the same. However, if the application is run by the original center, as was the case for MoodGYM and BluePages, the proliferation of different versions represents a logistical challenge, particularly if programs, such as the BluePages depression information website, requires updating as new information develops, or changes are introduced on the basis of user feedback. Fourthly, the contract should be explicit about the ownership of data collected from the translator's national users and trial. Finally, special attention should be paid to the timely updating of the translated version after changes in the original one.

Limitations

There are several limitations of this study. The validation part of this paper is based on mixed model repeated measures (MMRM) to handle missing data, which might have introduced some bias. However, the results used in this report were from the complete case analyses as we needed pretest and posttest for all cases to carry out the estimates for QALYs. The effect of the intention-to-treat sample was somewhat smaller than for the complete cases, and our choice of sample could have strengthened the effect and QOL calculations. The deterioration in the control group is not unusual over a short time [78], but over a longer time period, one would expect improvements. The effectiveness study discusses some plausible reasons for this [37], such as that the posttest was conducted just weeks before university examinations started. The usage of completers analysis could have introduced some bias, as well. The effects were considered in a relatively short time span of 8 weeks and should be regarded as potential effect if one completes the trial. There was no long-term follow-up of this trial. How the cost-effectiveness of Internet-based treatment is affected when a longer period of time is used is unknown to us. If an effect is maintained over a longer period, this will increase the QALY gain and thus increase the QALY and reduce the number needed to treat. This assumption was taken into the calculations based on the 12-month outcome data from other studies with MoodGYM and BluePages [63]. All outcome measures are based on self-reports and include only symptoms of depression because no clinical assessment was done. Self-report measures of depression may yield a high rate of false positive and false negative results. This could have compromised the results, identifying persons as depressed who do not meet the criteria for a clinical diagnosis and failing to identify persons with a depressive disorder. No data on sick leave, disability, or other social benefits were collected. The use of such data would have yielded a more complete and realistic cost-effectiveness estimate. However, these benefits would only have increased the estimated savings and present Internet-based interventions in an even more attractive way. Finally, the Rosser Index is just one way to estimate health gains, and other methods could have produced different results [79]. However, this method has been employed previously, and the Rosser Index has its strengths at a policy level often used on existing datasets to foretell health care demands [57]. A similar estimation for QALY gain was conducted to evaluate cost-effectiveness for a face-to-face group intervention designed to decrease depression [80], but they reported only QOL gain on a group level.

Conclusions

The economic analysis shows that the cost-effectiveness of the translation project was substantial. This is a natural consequence if the project is proven effective in reducing depressive symptoms and based on unguided self-help. The development cost is fixed and the maintenance cost is minimal. Further, as Internet-based treatments are more accessible than therapist-delivered treatments, when these Web-based interventions are disseminated at a national level, they should reach a significant number of individuals with an unmet need for help. The current findings should encourage others to undertake similar translation activities and facilitate global dissemination of effective programs to reduce health disparities across countries and cultures. When data on societal benefits are taken into account, Internet-based prevention interventions should become even more appealing.

Conflicts of Interest

Dr Griffiths is one of the authors and developers of the MoodGYM and BluePages websites, but derives neither personal nor financial benefit from the programs.

References

1. Marks IM, Cavanagh K, Gega L. Computer-aided psychotherapy: revolution or bubble? Br J Psychiatry 2007 Dec;191:471-473 [FREE Full text] [doi: 10.1192/bjp.bp.107.041152] [Medline: 18055948]
2. Freeman M, Pathare S. WHO resource book on mental health, human rights and legislation. Geneva, Switzerland: World Health Organization; 2005. URL: http://libdoc.who.int/publications/2005/924156282X.pdf [accessed 2012-10-15] [WebCite Cache ID 6BRV3KjAF]
3. World Health Organization. The global burden of disease: 2004 update. Geneva, Switzerland: World Health Organization; 2008. URL: http://www.who.int/healthinfo/global_burden_disease/GBD_report_2004update_full.pdf [accessed 2012-10-15] [WebCite Cache ID 6BRVD13mx]
4. Luppa M, Heinrich S, Angermeyer MC, König HH, Riedel-Heller SG. Cost-of-illness studies of depression: a systematic review. J Affect Disord 2007 Feb;98(1-2):29-43. [doi: 10.1016/j.jad.2006.07.017] [Medline: 16952399]
5. Katon W, Fan MY, Unützer J, Taylor J, Pincus H, Schoenbaum M. Depression and diabetes: a potentially lethal combination. J Gen Intern Med 2008 Oct;23(10):1571-1575 [FREE Full text] [doi: 10.1007/s11606-008-0731-9] [Medline: 18649108]
6. Smit F, Ederveen A, Cuijpers P, Deeg D, Beekman A. Opportunities for cost-effective prevention of late-life depression: an epidemiological approach. Arch Gen Psychiatry 2006 Mar;63(3):290-296. [doi: 10.1001/archpsyc.63.3.290] [Medline: 16520434]
7. Marks IM, Cavanagh K, Gega L. Hands-on Help: Computer-Aided Psychotherapy. Hove, UK: Psychology Press; 2007.
8. Clarke G, Eubanks D, Reid E, Kelleher C, O'Connor E, DeBar LL, et al. Overcoming Depression on the Internet (ODIN) (2): a randomized trial of a self-help depression skills program with reminders. J Med Internet Res 2005;7(2):e16 [FREE Full text] [doi: 10.2196/jmir.7.2.e16] [Medline: 15998607]
9. Christensen H, Griffiths KM, Jorm AF. Delivering interventions for depression by using the internet: randomised controlled trial. BMJ 2004 Jan 31;328(7434):265 [FREE Full text] [doi: 10.1136/bmj.37945.566632.EE] [Medline: 14742346]
10. Andersson G, Bergström J, Holländare F, Ekselius L, Carlbring P. Delivering cognitive behavioural therapy for mild to moderate depression via the Internet: predicting outcome at 6-month follow-up. Verhaltenstherapie 2004;14:185-189. [doi: 10.1159/000080914]

11. Griffiths KM, Christensen H. Review of randomised controlled trials of Internet interventions for mental disorders and related conditions. Clinical Psychologist 2006;10:16-29. [doi: 10.1080/13284200500378696]
12. Griffiths KM, Farrer L, Christensen H. InPsych. 2006. Effective treatments a click away URL: http://www.psychology.org.au/publications/inpsych/click/ [accessed 2013-01-20] [WebCite Cache ID 6DoeSDzPz]
13. Griffiths K, Farrer L, Christensen H. Clickety-click: e-mental health train on track. Australas Psychiatry 2007 Apr;15(2):100-108. [doi: 10.1080/10398560601123716] [Medline: 17464651]
14. Patten SB. Prevention of depressive symptoms through the use of distance technologies. Psychiatr Serv 2003 Mar;54(3):396-398. [Medline: 12610251]
15. van Straten A, Cuijpers P, Smits N. Effectiveness of a web-based self-help intervention for symptoms of depression, anxiety, and stress: randomized controlled trial. J Med Internet Res 2008;10(1):e7 [FREE Full text] [doi: 10.2196/jmir.954] [Medline: 18364344]
16. Spek V, Nyklícek I, Smits N, Cuijpers P, Riper H, Keyzer J, et al. Internet-based cognitive behavioural therapy for subthreshold depression in people over 50 years old: a randomized controlled clinical trial. Psychol Med 2007 Dec;37(12):1797-1806. [doi: 10.1017/S0033291707000542] [Medline: 17466110]
17. Meyer B, Berger T, Caspar F, Beevers CG, Andersson G, Weiss M. Effectiveness of a novel integrative online treatment for depression (Deprexis): randomized controlled trial. J Med Internet Res 2009;11(2):e15 [FREE Full text] [doi: 10.2196/jmir.1151] [Medline: 19632969]
18. Griffiths KM, Crisp D, Christensen H, Mackinnon AJ, Bennett K. The ANU WellBeing study: a protocol for a quasi-factorial randomised controlled trial of the effectiveness of an Internet support group and an automated Internet intervention for depression. BMC Psychiatry 2010;10:20 [FREE Full text] [doi: 10.1186/1471-244X-10-20] [Medline: 20211025]
19. Robertson L, Smith M, Castle D, Tannenbaum D. Using the Internet to enhance the treatment of depression. Australas Psychiatry 2006 Dec;14(4):413-417. [doi: 10.1111/j.1440-1665.2006.02315.x] [Medline: 17116083]
20. Ruwaard J, Schrieken B, Schrijver M, Broeksteeg J, Dekker J, Vermeulen H, et al. Standardized web-based cognitive behavioural therapy of mild to moderate depression: a randomized controlled trial with a long-term follow-up. Cogn Behav Ther 2009 Dec;38(4):206-221. [doi: 10.1080/16506070802408086] [Medline: 19221919]
21. Living Life to the Full. 2010. URL: http://www.llttf.com/index.php?section=page&page_seq=13 [accessed 2013-01-18] [WebCite Cache ID 6DlIj0SNh]
22. CCBT Limited Healthcare Online. 2010. MoodCalmer, online self help cCBT program for people experiencing low mood and depression URL: http://www.ccbt.co.uk/ [accessed 2012-10-15] [WebCite Cache ID 6BRUmmEGF]
23. Clark DM, Layard R, Smithies R, Richards DA, Suckling R, Wright B. Improving access to psychological therapy: initial evaluation of two UK demonstration sites. Behav Res Ther 2009 Nov;47(11):910-920 [FREE Full text] [doi: 10.1016/j.brat.2009.07.010] [Medline: 19647230]
24. Morgan AJ, Jorm AF, Mackinnon AJ. Email-based promotion of self-help for subthreshold depression: Mood Memos randomised controlled trial. Br J Psychiatry 2012 May;200(5):412-418. [doi: 10.1192/bjp.bp.111.101394] [Medline: 22442102]
25. Harrison V, Proudfoot J, Wee PP, Parker G, Pavlovic DH, Manicavasagar V. Mobile mental health: review of the emerging field and proof of concept study. J Ment Health 2011 Dec;20(6):509-524. [doi: 10.3109/09638237.2011.608746] [Medline: 21988230]
26. Van Voorhees BW, Fogel J, Reinecke MA, Gladstone T, Stuart S, Gollan J, et al. Randomized clinical trial of an Internet-based depression prevention program for adolescents (Project CATCH-IT) in primary care: 12-week outcomes. J Dev Behav Pediatr 2009 Feb;30(1):23-37. [doi: 10.1097/DBP.0b013e3181966c2a] [Medline: 19194326]
27. Titov N, Andrews G, Kemp A, Robinson E. Characteristics of adults with anxiety or depression treated at an internet clinic: comparison with a national survey and an outpatient clinic. PLoS One 2010;5(5):e10885 [FREE Full text] [doi: 10.1371/journal.pone.0010885] [Medline: 20526371]
28. Titov N, Dear BF, Schwencke G, Andrews G, Johnston L, Craske MG, et al. Transdiagnostic internet treatment for anxiety and depression: a randomised controlled trial. Behav Res Ther 2011 Aug;49(8):441-452. [doi: 10.1016/j.brat.2011.03.007] [Medline: 21679925]
29. McKeown A, Potts HWW. Computerised cognitive-behavioural therapy for prevention early intervention in anxiety and depression: a case study of Xanthis. 2009. URL: http://discovery.ucl.ac.uk/18338/1/18338.pdf [accessed 2013-01-20] [WebCite Cache ID 6Dof83TqX]
30. Deitz DK, Cook RF, Billings DW, Hendrickson A. A web-based mental health program: reaching parents at work. J Pediatr Psychol 2009 Jun;34(5):488-494 [FREE Full text] [doi: 10.1093/jpepsy/jsn108] [Medline: 18845585]
31. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. Cogn Behav Ther 2009 Dec;38(4):196-205. [doi: 10.1080/16506070903318960] [Medline: 20183695]
32. Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis. PLoS One 2010;5(10):e13196 [FREE Full text] [doi: 10.1371/journal.pone.0013196] [Medline: 20967242]
33. World Health Organization, Department of Mental Health and Substance Abuse, Prevention Research Centre of the Universities of Nijmegen and Maastricht. Prevention of mental disorders: effective interventions and policy options:

summary report. Geneva, Switzerland: World Health Organization; 2004. URL: http://www.who.int/mental_health/evidence/en/prevention_of_mental_disorders_sr.pdf [accessed 2012-10-15] [WebCite Cache ID 6BRU5qBZD]
34. 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]
35. Marks I, Cavanagh K. Computer-aided psychological treatments: evolving issues. Annu Rev Clin Psychol 2009;5:121-141. [doi: 10.1146/annurev.clinpsy.032408.153538] [Medline: 19327027]
36. Titov N. Internet-delivered psychotherapy for depression in adults. Curr Opin Psychiatry 2011 Jan;24(1):18-23. [doi: 10.1097/YCO.0b013e32833ed18f] [Medline: 20827199]
37. Lintvedt OK, Griffiths KM, Sørensen K, Ostvik AR, Wang CE, Eisemann M, et al. Evaluating the effectiveness and efficacy of unguided internet-based self-help intervention for the prevention of depression: a randomized controlled trial. Clin Psychol Psychother 2013 Jan;20(1):10-27. [doi: 10.1002/cpp.770] [Medline: 21887811]
38. Carlton PA, Deane FP. Impact of attitudes and suicidal ideation on adolescents' intentions to seek professional psychological help. J Adolesc 2000 Feb;23(1):35-45. [doi: 10.1006/jado.1999.0299] [Medline: 10700370]
39. Calear AL, Christensen H, Mackinnon A, Griffiths KM, O'Kearney R. The YouthMood Project: a cluster randomized controlled trial of an online cognitive behavioral program with adolescents. J Consult Clin Psychol 2009 Dec;77(6):1021-1032. [doi: 10.1037/a0017391] [Medline: 19968379]
40. Ritterband LM, Thorndike FP, Gonder-Frederick LA, Magee JC, Bailey ET, Saylor DK, et al. Efficacy of an Internet-based behavioral intervention for adults with insomnia. Arch Gen Psychiatry 2009 Jul;66(7):692-698. [doi: 10.1001/archgenpsychiatry.2009.66] [Medline: 19581560]
41. Spek V, Cuijpers P, Nyklícek I, Riper H, Keyzer J, Pop V. Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. Psychol Med 2007 Mar;37(3):319-328. [doi: 10.1017/S0033291706008944] [Medline: 17112400]
42. Ritterband LM, Andersson G, Christensen HM, Carlbring P, Cuijpers P. Directions for the International Society for Research on Internet Interventions (ISRII). J Med Internet Res 2006;8(3):e23 [FREE Full text] [doi: 10.2196/jmir.8.3.e23] [Medline: 17032639]
43. Australian National University. The MoodGYM training program. 2012. URL: https://www.moodgym.anu.edu.au/welcome/faq [accessed 2012-10-15] [WebCite Cache ID 6BRTz2zMg]
44. Australian National University. BluePages Depression Information. 2012. URL: http://www.bluepages.anu.edu.au/about_bluepages/ [accessed 2012-10-15] [WebCite Cache ID 6BRTeN0iB]
45. Bennett K, Reynolds J, Christensen H, Griffiths KM. e-hub: an online self-help mental health service in the community. Med J Aust 2010 Jun 7;192(11 Suppl):S48-S52. [Medline: 20528710]
46. Tate DF, Finkelstein EA, Khavjou O, Gustafson A. Cost effectiveness of internet interventions: review and recommendations. Ann Behav Med 2009 Aug;38(1):40-45 [FREE Full text] [doi: 10.1007/s12160-009-9131-6] [Medline: 19834778]
47. Lintvedt OK, Sørensen K, Østvik AR, Verplanken B, Wang CE. The need for web-based cognitive behavior therapy among university students. Journal of Technology in Human Services 2008;26(2-4):239-258. [doi: 10.1080/15228830802096705]
48. Statistics Norway. 2012. Consumer Price Index URL: http://www.ssb.no/vis/kpi/kpiregn.html [accessed 2012-10-15] [WebCite Cache ID 6BRTOqybE]
49. Reserve Bank of Australia. 2012. Inflation calculator URL: http://www.rba.gov.au/calculator/annualDecimal.html [accessed 2012-10-15] [WebCite Cache ID 6BRSzRuMA]
50. Radloff LS. The CES-D scale: A self-report depression scale for research in the general population. Applied Psychological Measurement 1977 Jun;1(3):385-401. [doi: 10.1177/014662167700100306]
51. Santor D, Zuroff DC, Ramsay JO, Cervantes P, Palacios J. Examining scale discriminability in the BDI and CES-D as a function of depressive severity. Psychological Assessment 1995 Jun;7(2):131-139. [doi: 10.1037/1040-3590.7.2.131]
52. Rushton JL, Forcier M, Schectman RM. Epidemiology of depressive symptoms in the National Longitudinal Study of Adolescent Health. J Am Acad Child Adolesc Psychiatry 2002 Feb;41(2):199-205. [doi: 10.1097/00004583-200202000-00014] [Medline: 11837410]
53. Kessler RC, Andrews G, Colpe LJ, Hiripi E, Mroczek DK, Normand SL, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. Psychol Med 2002 Aug;32(6):959-976. [Medline: 12214795]
54. Gater RA, Kind P, Gudex C. Quality of life in liaison psychiatry. A comparison of patient and clinician assessment. Br J Psychiatry 1995 Apr;166(4):515-520. [Medline: 7795926]
55. McCabe C. What is cost-utility analysis?. London: Hayward Medical Communications; 2009. URL: http://www.medicine.ox.ac.uk/bandolier/painres/download/whatis/What_is_cost-util.pdf [accessed 2012-10-15] [WebCite Cache ID 6BRW35q5E]
56. Rosser R, Kind P. A scale of valuations of states of illness: is there a social consensus? Int J Epidemiol 1978 Dec;7(4):347-358. [Medline: 744673]
57. Hawthorne G, Richardson J, Olsen JA, Mortimer D, Smith R. Critical review of the multi-attribute utility literature. West Heidelberg, Australia: Centre for Health Program Evaluation, Monash University; 2001. URL: http://www.buseco.monash.edu.au/centres/che/pubs/wp103.pdf [accessed 2012-10-15] [WebCite Cache ID 6BRWJr9Vc]

58. Gudex C, Kind P, van Dalen H, Durand MA, Morris J, Williams A. Comparing scaling methods for health state valuations: Rosser revisited (Discussion paper 107). York, UK: Centre for Health Economics, University of York; 1993. URL: http://www.york.ac.uk/media/che/documents/papers/discussionpapers/CHE%20Discussion%20Paper%20107.pdf [accessed 2012-10-15] [WebCite Cache ID 6BRX9FI4j]
59. Kind P, Rosser R, Williams A. Valuation of quality of life: some psychometric evidence. In: Jones-Lee MW, editor. The Value of Life and Safety. Amsterdam: North-Holland Pub Co; 1982:159-170.
60. Nord E. [A table of reference values for cost-benefit-analyses in health care]. Tidsskr Nor Laegeforen 1996 Nov 10;116(27):3246-3249. [Medline: 9011980]
61. Gudex C, Kind P. The QALY toolkit (Discussion paper 38). York, UK: University of York Centre for Health Economics; 1988. URL: http://www.york.ac.uk/media/che/documents/papers/discussionpapers/CHE%20Discussion%20Paper%2038.pdf [accessed 2012-10-15] [WebCite Cache ID 6BRXdpg2u]
62. Torrance GW, Feeny D. Utilities and quality-adjusted life years. Int J Technol Assess Health Care 1989;5(4):559-575. [Medline: 2634630]
63. Mackinnon A, Griffiths KM, Christensen H. Comparative randomised trial of online cognitive-behavioural therapy and an information website for depression: 12-month outcomes. Br J Psychiatry 2008 Feb;192(2):130-134 [FREE Full text] [doi: 10.1192/bjp.bp.106.032078] [Medline: 18245031]
64. Olsen JA. In: Olsen J, editor. Principles in Health Economics and Policy: Distributing Health Care. Oxford, UK: Oxford University Press; 2009.
65. National Institute for Health and Clinical Excellence. NICE Guideline Manual. 2007. Incorporating health economics in guidelines and assessing resource impact URL: http://www.nice.org.uk/niceMedia/pdf/GuidelinesManualChapter8.pdf [accessed 2012-10-15] [WebCite Cache ID 6BRSXCC7y]
66. National Institute for Health and Clinical Excellence. 2010. Measuring effectiveness and cost effectiveness: the QALY URL: http://www.nice.org.uk/newsroom/features/measuringeffectivenessandcosteffectivenesstheqaly.jsp [accessed 2012-10-15] [WebCite Cache ID 6BRSOCbIr]
67. Towse A, Pritchard C. National Institute for Clinical Excellence (NICE): Is economic appraisal working? Pharmacoeconomics 2002;20 Suppl 3:95-105. [Medline: 12457430]
68. Norwegian Directorate for Health and Social Affairs. Helseeffekter i samfunnsøkonomiske analyser [Health effects of economic analyses]. 2007. URL: http://www.shdir.no/vp/multimedia/archive/00020/IS-1435_20969a.pdf [accessed 2012-10-15] [WebCite Cache ID 6BRSFua0o]
69. Butler J, Christensen H, Bartlett LE, Groves C. Cost-effectiveness analysis of MoodGYM in the management of clinical depression. 2005. URL: http://www.acerh.edu.au/Materials/WPA_MoodGYM_Abstract.pdf [accessed 2012-10-15] [WebCite Cache ID 6BRRtfERu]
70. Cuijpers P, van Straten A, Smit F, Mihalopoulos C, Beekman A. Preventing the onset of depressive disorders: a meta-analytic review of psychological interventions. Am J Psychiatry 2008 Oct;165(10):1272-1280. [doi: 10.1176/appi.ajp.2008.07091422] [Medline: 18765483]
71. Botella C, Hofmann SG, Moscovitch DA. A self-applied, Internet-based intervention for fear of public speaking. J Clin Psychol 2004 Aug;60(8):821-830. [doi: 10.1002/jclp.20040] [Medline: 15241810]
72. 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]
73. Lange A, van de Ven JP, Schrieken B. Interapy: treatment of post-traumatic stress via the internet. Cogn Behav Ther 2003;32(3):110-124. [doi: 10.1080/16506070302317] [Medline: 16291543]
74. Richards J, Klein B, Carlbring P. Internet-based treatment for panic disorder. Cogn Behav Ther 2003;32(3):125-135. [doi: 10.1080/16506070302318] [Medline: 16291544]
75. Winzelberg AJ, Eppstein D, Eldredge KL, Wilfley D, Dasmahapatra R, Dev P, et al. Effectiveness of an Internet-based program for reducing risk factors for eating disorders. J Consult Clin Psychol 2000 Apr;68(2):346-350. [Medline: 10780136]
76. Reynolds CF. Prevention of depressive disorders: a brave new world. Depress Anxiety 2009;26(12):1062-1065 [FREE Full text] [doi: 10.1002/da.20644] [Medline: 19957277]
77. Warmerdam L, Smit F, van Straten A, Riper H, Cuijpers P. Cost-utility and cost-effectiveness of internet-based treatment for adults with depressive symptoms: randomized trial. J Med Internet Res 2010;12(5):e53 [FREE Full text] [doi: 10.2196/jmir.1436] [Medline: 21169166]
78. Titov N, Andrews G, Choi I, Schwencke G, Mahoney A. Shyness 3: randomized controlled trial of guided versus unguided Internet-based CBT for social phobia. Aust N Z J Psychiatry 2008 Dec;42(12):1030-1040. [doi: 10.1080/00048670802512107] [Medline: 19016091]
79. Saxena S, Carlson D, Billington R, WHOQOL Group. World Health Organisation Quality Of Life. The WHO quality of life assessment instrument (WHOQOL-Bref): the importance of its items for cross-cultural research. Qual Life Res 2001;10(8):711-721. [Medline: 11871592]
80. Nord E, Dalgard OS. [Economic evaluation of a course in coping with depression]. Tidsskr Nor Laegeforen 2006 Feb 23;126(5):586-588 [FREE Full text] [Medline: 16505865]

Abbreviations

ANU: Australian National University
CBT: cognitive behavior therapy
CER: cost-effectiveness ratio
CES-D: Center for Epidemiologic Studies Depression Scale
d: Cohen’s d
g: Hedges’ g
ΔH: degree of improvement in health
ΔHext: degree of improvement in health, extrapolated
ISRII: International Society for Research on Internet Interventions
K10: Kessler Psychological Distress Scale
MMRM: mixed model repeated measures
NICE: National Institute for Clinical Excellence
NOK: Norwegian kroner
QALY: quality-adjusted life year
QALY-gain: QALYs gained over a timeframe of 1 year
QOL: quality of life
RCT: randomized controlled trial
WHO: World Health Organization

Edited by G Eysenbach; submitted 05.11.12; peer-reviewed by S Maloney, D Kavanagh; comments to author 27.11.12; revised version received 03.12.12; accepted 20.12.12; published 23.01.13.

Please cite as:
Lintvedt OK, Griffiths KM, Eisemann M, Waterloo K
Evaluating the Translation Process of an Internet-Based Self-Help Intervention for Prevention of Depression: A Cost-Effectiveness Analysis
J Med Internet Res 2013;15(1):e18
URL: http://www.jmir.org/2013/1/e18/
doi:10.2196/jmir.2422
PMID:23343481

Original Paper

Engagement Promotes Abstinence in a Web-based Cessation Intervention: Cohort Study

Amanda Richardson[1,2], MS, PhD; Amanda L Graham[3,4], PhD; Nathan Cobb[3,4], MD; Haijun Xiao[1], MS; Aaron Mushro[5], MS, MBA; David Abrams[2,3,4], PhD; Donna Vallone[1,2], MPH, PhD

[1]Department of Research and Evaluation, Legacy, Washington, DC, United States

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

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

[4]Lombardi Comprehensive Cancer Center, Georgetown University Medical Center, Washington, DC, United States

[5]Department of Marketing, Legacy, Washington, DC, United States

Corresponding Author:
Amanda Richardson, MS, PhD
Department of Research and Evaluation
Legacy
1724 Massachusetts Avenue NW
Washington, DC, 20036
United States
Phone: 1 202 454 5571
Fax: 1 202 454 5599
Email: arichardson@legacyforhealth.org

Abstract

Background: Web-based smoking cessation interventions can have a public health impact because they are both effective in promoting cessation and can reach large numbers of smokers in a cost-efficient manner. Their potential impact, however, has not been realized. It is still unclear how such interventions promote cessation, who benefits most, and how to improve their population impact.

Objective: To examine the effectiveness of a highly promoted Web-based smoking cessation intervention to promote quit behavior over time, identify the most effective features, and understand who is most likely to use those features by using unweighted and weighted analyses to estimate the impact in the broader pool of registered site users.

Methods: A sample of 1033 new adult registrants was recruited from a Web-based smoking cessation intervention by using an automated study management system. Abstinence was assessed by self-report through a mixed-mode follow-up (online survey with telephone follow-up for nonrespondents) at 1, 3, and 6 months. Software tracked respondents' online activity. Generalized estimating equations (GEE) were used to examine predictors of website utilization and how utilization promoted abstinence using unweighted and weighted data.

Results: The 7-day point prevalence abstinence rates at 6 months ranged from 20.68% to 11.13% in the responder and intent-to-treat samples, respectively. Predictors of abstinence in unweighted analyses included number of visits to the website as well as accessing specific interactive or engaging features. In weighted analyses, only number of visits was predictive of abstinence. Motivation to quit was a key predictor of website utilization, whereas negative partner support decreased the likelihood of increasing visits or accessing engaging features.

Conclusions: Engagement is critical to promoting smoking cessation. The next generation of Web-based smoking cessation interventions needs to maximize the initial engagement of all new visitors and work to retain those smokers who proceed to register on the site.

(J Med Internet Res 2013;15(1):e14) doi:10.2196/jmir.2277

KEYWORDS

Internet; smoking cessation; intervention studies; use effectiveness

Introduction

Despite a substantial decline in smoking prevalence over the past 40 years, nearly 1 in 5 adults (45.3 million) in the United States still smokes, and the rate of prevalence reduction has slowed in the last decade [1]. Most smokers (approximately 70%) say that they want to quit, and each year roughly one-half of smokers try to quit, but less than 6% succeed [2]. Furthermore, only 31.7% have used counseling and/or medications when they tried to quit [2], highlighting the need for cost-effective, engaging, and sustainable cessation interventions that can reach and motivate smokers to quit and help them remain abstinent. Given the substantial influence of the Web, with its over 2.67 billion users across the globe [3] and its increasing use as a mechanism to seek health information [4,5], Web-based smoking cessation interventions show significant potential to promote cessation if optimized to be maximally effective (a product of reach and efficacy divided by cost) [6,7].

The 2008 Clinical Practice Guidelines for Treating Tobacco Dependence [8] assert that Web-based interventions are a “highly promising delivery system for tobacco dependence treatment.” These interventions can be built to run on multiple platforms including tablets and mobile devices [9-11], are easily accessible and available all hours of the day, provide anonymity to those smokers who prefer privacy [12], and remove the logistical and financial barriers associated with some pharmaceutical and/or in-person behavioral counseling interventions. Additionally, Web-based interventions have evolved to engage smokers with interactive elements including multimedia content, social networks, blogs, forums, and virtual communities [9,13-15]. Web-based interventions show evidence of effectiveness in adult tobacco users [16-18], with recent randomized trials reporting quit rates as high as 26% at 6 months posttreatment [19-21]. Although several studies have documented increased rates of cessation associated with tailoring [14,22-27], greater website exposure [25,28-36], and a supportive vs information-only content [24,33,34], a recent review concluded that the strength of the evidence for effectiveness is only moderate and could benefit from additional research to improve population impact [37]. This review also concluded that there is currently insufficient evidence for the effectiveness of multicomponent, interactive, or additional components, such as bulletin boards, in promoting cessation [37]. Consistent with the National Institutes of Health’s strategic priorities, there is an extraordinary opportunity to conduct more rigorous implementation/dissemination research to improve the real-world impact of Web-based interventions for smoking cessation [38-40].

Tobacco control simulation models suggest that Web-based interventions have extraordinary potential to contribute to a further decrease in population-level smoking rates if effective and if promoted sufficiently [41]. For the full potential of these interventions to be reached, however, further research is needed to clarify the degree to which Web-based interventions are effective, how and why these interventions work, the factors influencing their effectiveness, and their long-term benefits in a real-world setting.

The purpose of this study was to build upon previous literature to examine the effectiveness of an interactive Web-based cessation intervention, BecomeAnEX.org [42], in promoting quit attempts and abstinence over a 6-month period in a large sample of newly registered users. Specific aims of the study were to (1) identify whether increased use of the website promotes quit attempts and abstinence, (2) determine which components of the program are most effective in promoting cessation, and (3) examine the demographic characteristics, psychosocial factors, and smoking behaviors associated with greater use of the program and its associated features. We hypothesized that increased utilization of the website and accessing the BecomeAnEX.org community would be associated with quit behavior over the course of this study.

Methods

The BecomeAnEX.org Intervention

BecomeAnEX.org is a free, branded, evidence-based intervention developed in accordance with the 2008 US Department of Health and Human Service’s Clinical Practice Guidelines [8]. The site was developed by Legacy, a nonprofit organization that develops smoking prevention and cessation programs, in collaboration with the Mayo Clinic Nicotine Dependence Center. The overarching goal of the site is to educate smokers wanting to quit and provide the tools necessary to support their quitting efforts. The site provides smokers with detailed information and action steps to help them “re-learn life without cigarettes,” in part by disassociating smoking from common daily activities that would otherwise function as smoking cues, such as driving or drinking coffee. BecomeAnEX.org is designed to support smokers through their journey of quitting smoking and preventing relapse. BecomeAnEX.org is available on multiple platforms, including mobile, and is designed to engage smokers through the following core elements: (1) My Quit Plan, a checklist that helps smokers set a quit date and tracks their progress through the steps of the quit plan; (2) a Cigarette Tracker to help smokers identify and track their smoking triggers; (3) Beat Your Smoking Triggers, a list of common triggers with tips for dissociating cigarettes from them; (4) separation exercises to teach smokers how to separate a cigarette from its trigger; (5) support exercises to help smokers identify who they can turn to for support when quitting smoking; (6) Quit Smoking Resources, a list of national organizations, state quitlines, and websites that smokers can turn to for additional information and help to quit smoking; (7) text and videos detailing the importance of medication and the different types available; and (8) BecomeAnEX.org Community, a large online network of current and former smokers. See Figures 1 and 2 for screenshots of the website.

Figure 1. Screenshot of the home page of the BecomeAnEX.org website.

A New Way To Think About Quitting

The **EX** Plan is a free quit smoking program that helps you re-learn your life without cigarettes. Before you actually stop smoking, we'll show you how to deal with the very things that trip up so many people when they try to quit smoking. So you'll be more prepared to quit and stay quit.

Whether this is your first try at quitting smoking or your 10th, the free **EX** Plan can help you really do it this time. LEARN MORE ABOUT THE PLAN.

WATCH A VIDEO on what a real ex-smoker has to say about EX

The EX Plan iPhone™ App

Get the best of BecomeAnEX and re-learning life without cigarettes in an iPhone™ app. Introducing the EX Plan app, a whole new way to think about quitting smoking.

BecomeAnEX.org is brought to you by Legacy to help people quit smoking.
BecomeAn**EX** is only responsible for what we post, not what others post or what's linked to us in the community section.
© 2012 American Legacy Foundation

Figure 2. Screenshot of the My Quit Plan page of the BecomeAnEX.org website.

Register now for free, so you can personalize your own "My Quit Plan homepage," access tools like the Cigarette Tracker and the Quit Date Calendar, and track your progress with the Quit Smoking Checklist.

With over 300,000 members and counting, you'll never be alone in your mission to quit.

You've made it through 0% of your quit plan.

Setting a quit date and tracking your cigarettes are both important parts of the EX Plan. Use the calendar below to select your quit date.

Set or change your quit date by clicking on a day below.

thebestbattle posted on Owlfeather's message board

Linda N posted on pir8fan's message board

I can't remember the day lol

So sometimes my boss can be such a doukjhnfe. But it...

Don't step on any toes.

I really admire some of you here that...

The Community feels your pain.

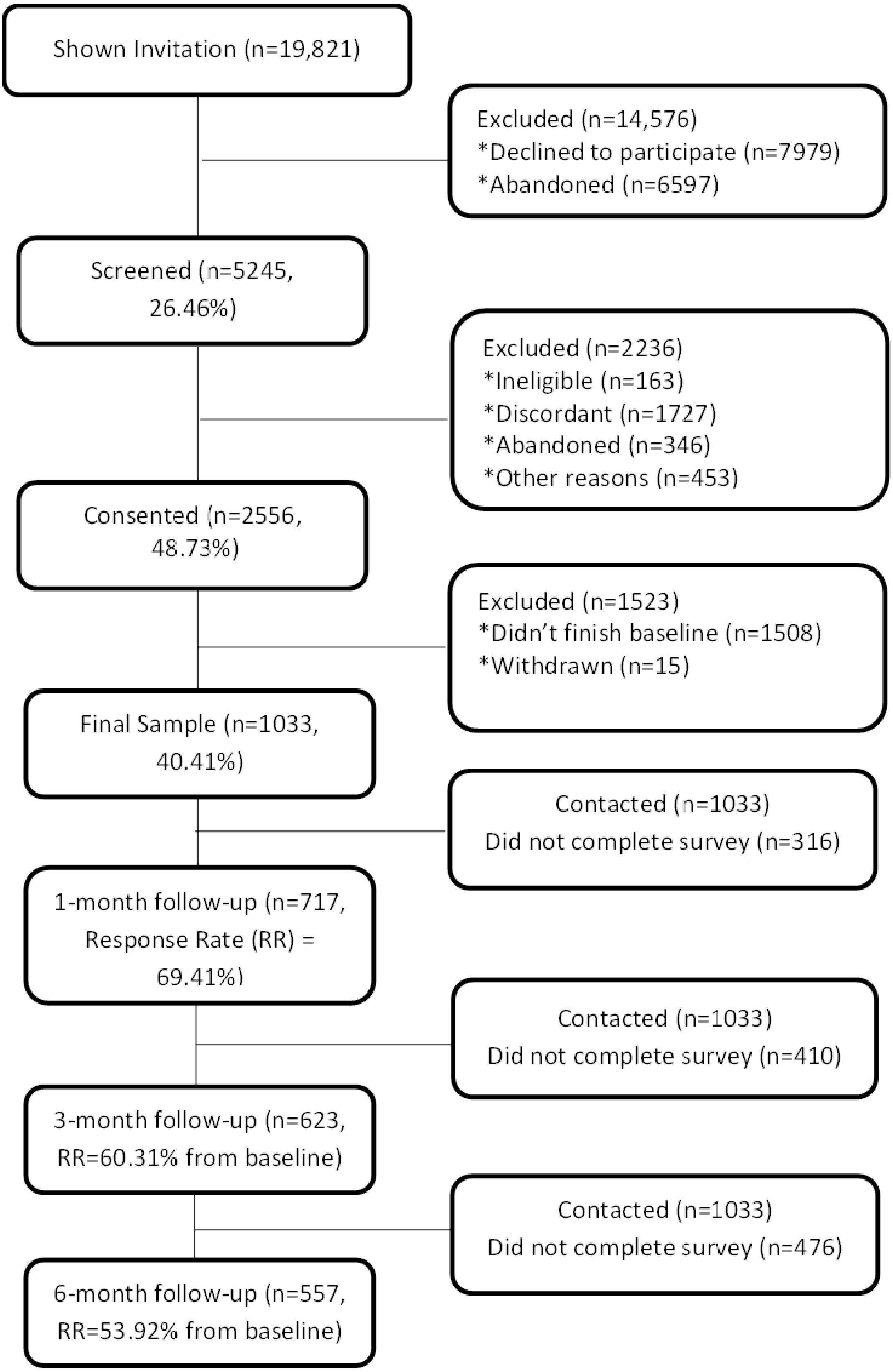
Nobody gets what you're going through like the EX Community. Get the support you need to successfully quit.

BecomeAnEX.org is brought to you by Legacy to help people quit smoking.
BecomeAnEX is only responsible for what we post, not what others post or what's linked to us in the community section.
© 2012 American Legacy Foundation

Study Participants and Enrollment

This was a prospective, observational study of a cohort of participants recruited via an automated study management system at the time of enrollment on BecomeAnEX.org. To be eligible, BecomeAnEX.org registrants had to be current smokers, age 18 years or older, and willing to provide a valid email address and telephone number. From January 19 through May 23, 2011, a total of 19,821 individuals were invited to participate in the study (Figure 3). Enrollment for certain demographic subgroups (eg, white females) was deactivated once targeted enrollment goals were met. The invitation described the terms of the study, expectations regarding follow-up assessments, and the availability of incentives for completion of follow-up surveys. Registrants who accepted the invitation (n=5245) completed online eligibility verification and online informed consent. Those who were eligible (current smokers, age ≥18 years, provided valid email/telephone number) and consented (n=2556) were sent an email with a link to the baseline survey that they had to complete within 24 hours. A total of 1048 individuals completed the baseline survey; of those, 15 were subsequently withdrawn because of duplication in study registration, leaving an effective sample size of 1033. Respondents were then contacted for follow-up surveys at 1, 3, and 6 months after enrollment. Intensive email and telephone follow-up yielded response rates of 69.41% (717/1033), 60.31% (623/1033), and 53.92% (557/1033), respectively.

Figure 3. CONSORT diagram of study recruitment.



Data Collection

Data for the study were obtained through 2 sources: (1) self-report assessments completed at baseline and at 1, 3, and 6 months after enrollment, and (2) online tracking software. At each follow-up, respondents were sent an email offering an incentive of US $25 to complete an online follow-up survey. Reminder emails were sent to participants 3 and 7 days after the initial email. Attempts were made to contact nonrespondents by phone to complete the follow-up survey starting at 10 days after the initial email. At least 15 but no more than 20 attempts were made to contact nonrespondents over multiple days at different times of day. Participants were reimbursed US $15 for completing the follow-up assessment by phone. The lower reimbursement offered through the telephone follow-up was intended to encourage online follow-up. Individual-level tracking metrics of BecomeAnEX.org utilization (eg, number of log-ins, pages viewed) were recorded by using Omniture SiteCatalyst software [43] and transferred to a relational database linked to the automated study management system. Utilization of BecomeAnEX.org via mobile devices was not recorded. At the time of the study, the mobile site associated with BecomeAnEx.org was a limited version that contained only a portion of the functionality of the full site and represented only anonymous use (ie, there was no log-in required or available). As such, although use of the mobile site was tracked by Omniture SiteCatalyst, it was not possible to record individual use by study participants. Websites other than BecomeAnEX.org were not monitored. This study received Institutional Review Board (IRB) approval by Schulman Associates IRB (previously Independent IRB, Inc).

Weighting

Data obtained on the study sample were poststratified to be representative of the larger population of new registrants to BecomeAnEX.org during the period of study recruitment and adjusted for nonresponse. Variables on which the study sample were weighted include gender, age, education, and race/ethnicity of new registrants. Each of these demographic variables was significantly correlated ($P<.05$) with baseline levels of the smoking variables mentioned previously (eg, motivation to quit, quit attempts in past year, and cigarettes per day). Demographic information was also significantly correlated with website usage during the study period.

Measures

The main outcomes were self-reported number of quit attempts, 7-day abstinence, and 30-day abstinence measured at each of the 3 follow-ups. Self-report of smoking behavior is commonly accepted in national Web-based cessation trials [19,20,22,23,29,30,44] when biochemical verification is neither feasible nor necessary [45]. Quit attempts were assessed with the question: "Since our last survey, how many times have you intentionally quit smoking (not even a puff) for at least 24 hours?" Abstinence at 7 and 30 days was measured at each follow-up with the question: "About how long has it been since you completely stopped smoking cigarettes?"

Main independent variables to measure website "dose" were 3 Omniture SiteCatalyst metrics recorded at each time period: (1) the number of visits to BecomeAnEX.org, (2) the number of minutes on the website, and (3) the number of pages accessed on the website. All exposure metrics recorded represent unmodified industry standards preset in Omniture SiteCatalyst. Specifically, time online was calculated as the duration between the first page view and last page view of a session, whereas a session times out after 30 minutes of inactivity. Cumulative values for each of these metrics, as measured by Omniture SiteCatalyst, were recorded at each follow-up survey. Additionally, data on the features accessed by the respondents (detailed in the Intervention section) were collected and examined. The top 5 most commonly accessed BecomeAnEX.org website features or pages within the study sample were examined in relation to quit behavior, namely: (1) My Quit Plan, (2) the BecomeAnEX.org Community, (3) separation exercises, (4) Beat Your Smoking Triggers, and (5) the "set-a-quit-date" feature. Usage of these features was coded categorically with 3 levels: never accessed, accessed once, and accessed 2 or more times. Additional metrics of interest were self-reported use of the BecomeAnEX.org Facebook and Twitter accounts that encouraged use of and provided links to the full site.

Control variables included age, race/ethnicity (non-Hispanic white, non-Hispanic black, Hispanic, and other), and education, which was categorized as less than a high school diploma, high school diploma or GED (General Educational Development) test, 1-3 years of college, and college degree (or higher). Baseline smoking variables that may potentially be related to both website utilization and quit behavior were also measured, including motivation to quit as measured by the question: "On a scale of 1 to 10, how much do you want to quit smoking?", number of quit attempts in the past year, nicotine dependence as assessed by the Fagerstrom Test for Nicotine Dependence [46], peer smoking (as measured by the question "Think about the 5 people outside of your family that you spend the most time with. How many of them smoke?"), household smoking (yes/no), partner smoking status (yes/no), and past year use of at least 1 cessation aide (eg, pharmacotherapy). The following psychosocial variables were measured at baseline given their association to quit behavior: social support as measured by a modified 6-item version of the Partner Interaction Questionnaire (PIQ) [47], health status (excellent, very good, good, fair, poor), advice to quit from a health care provider (yes/no), sadness as measured by the 12-item Short Form Health Survey (SF-12) [48] on a 6-point scale (with the question "How much of the time during the past 4 weeks have you felt downhearted and blue?"), perceived stress as measured by Cohen's Perceived Stress Scale [49], and frequency of Internet use.

Analysis

Data on the study cohort were analyzed in comparison to all other new registrants on the website during the period of study recruitment to assess differences between the study population and the more general pool of registrants. Registration data included demographic information and several questions assessing smoking behaviors, including cigarettes per day, motivation to quit, and quit attempts in the past year.

Both unweighted and weighted analyses of primary outcomes were completed. First, frequencies and descriptive data were calculated for baseline demographic characteristics, smoking behaviors, and psychosocial factors. Next, the median and interquartile range (IQR) of each of the utilization metrics and the frequency of utilization of each of the BecomeAnEX.org features were calculated. Smoking behavior at each follow-up survey was calculated by using self-reported smoking status. Abstinence rates were calculated 2 ways: (1) responder-only analyses that only considered participants at each follow-up with complete data, and (2) intent-to-treat (ITT) analyses that considered all study participants and counted participants lost to follow-up as smokers. Abstinence was also calculated as a function of the number of visits to the website.

Logistic regression analyses by using generalized estimating equations (GEE) were used to examine the association over time between website use and quit behavior [50]. Briefly, GEE accounts for correlations in repeated measurements and produces efficient regression parameter estimates and accurate standard errors [51]. Two models were calculated. The first model estimated the association between website usage variables (eg, number of visits) and quit behavior, adjusting for potentially important confounding variables mentioned in the Measures section. Because of the skew of the website usage variables, each was log-transformed. Since visits to the website were highly correlated with number of minutes on the website (r=0.84, P<.001) and number of pages accessed on the website (r=0.78, P<.001), results are presented only for number of visits. The second model adds in 4 of the top 5 Web features or pages accessed by participants and coded usage categorically into 3 levels: never accessed, accessed once, and accessed 2 or more times. Although My Quit Plan was the most-accessed page, it was removed from the analysis because of the low variability in its use (accessed by approximately 90% of the sample) and because it encompasses and provides links to the other features of the site. An additional set of models examined predictors of number of visits to the website as well as usage of the 2 features of BecomeAnEX.org that were significantly associated with abstinence, namely the BecomeAnEX.org community and separation exercises. These models included all the control variables described previously.

Results

Unweighted Data

The demographic characteristics of study participants are displayed in Table 1. There was a roughly even split between males (48.31%, 499/1033) and females (51.69%, 534/1033) and most participants were 25 to 49 years of age (85.38%, 882/1033). Consistent with our targeted recruitment goals, there was good representation of non-Hispanic blacks (14.04%, 145/1033) and Hispanics (14.52%, 150/1033), which closely mirrors that of the US population (13.1% and 16.7%, respectively) [52]. The sample was highly motivated to quit (mean 8.5 out of a maximum of 10, where 10 equals the highest level of motivation), and had made an average of 3 quit attempts in the prior year.

Table 1. Baseline characteristics of study participants using the BecomeAnEX.org smoking cessation website (N=1033).

Demographic variables	n (%)	Mean (SE)
Gender		
Female	534 (51.69)	
Male	499 (48.31)	
Race/ethnicity		
Non-Hispanic white	700 (67.76)	
Non-Hispanic black	145 (14.04)	
Hispanic	150 (14.52)	
Other	38 (3.68)	
Age, years		
18-24	120 (11.62)	
25-44	531 (51.40)	
45-64	351 (33.98)	
≥65	31 (3.00)	
Education		
Less than high school	62 (6.00)	
High school diploma/GED	240 (23.23)	
Some College	483 (46.76)	
College degree or higher	248 (24.01)	
Smoking variables		
Fagerstrom Test for Nicotine Dependence[a]		5.22 (0.07)
Motivation to quit (0-10)		8.48 (0.07)
Number of quit attempts in the past year		3.06 (0.20)
Number of smoking peers (0-5)		2.05 (0.05)
Has a partner who smokes	332 (32.14)	
Household smoking	581 (56.24)	
Advised to quit by a health care professional	847 (82.00)	
Health status		
Excellent	72 (6.97)	
Very good	304 (29.43)	
Good	396 (38.33)	
Fair	205 (19.85)	
Poor	56 (5.42)	
Psychological status		
Baseline sadness (1-6)[b]		3.45 (0.04)
Cohen Perceived Stress Scale, baseline		7.51 (0.09)
Baseline partner support [c]		
Positive support		11.28 (0.10)
Negative support		9.00 (0.12)
Frequency of use of the Internet		
Several times a day	807 (78.12)	
Approximately once a day	162 (15.68)	

Demographic variables	n (%)	Mean (SE)
<5 days a week	64 (6.20)	

[a] Range of responses for Fagerstrom Test was was 0-10.

[b] Measured using the 12-Item Short Form Health Survey (SF-12) questionnaire.

[c] Measured using the 6-item modified Partner Interaction Questionnaire (PIQ).

A comparison of study participants to other BecomeAnEX.org members that joined the site during the study period showed that study participants were more likely than non-study participants to smoke at least a pack of cigarettes a day (38.88% vs 25.57%, *P*<.001), more motivated to quit within the next 30 days (75.24% vs 66.10%, *P*<.001), and more likely to have made at least 1 quit attempt in the past year (64.88% vs 50.43%, *P*<.001) [data not shown].

Website Utilization

The median for total website visits over the 6-month study period was 2 (IQR 1-4), the median number of total minutes on BecomeAnEX.org was 32 (IQR 17-57), and the median number of total pages accessed was 24 (IQR 12-60). Of the total sample, 40.27% (416/1033) visited the website once, 44.63% (461/1033) visited 2 to 5 times, and the remaining 15.10% (156/1033) visited more than 5 times. The quit plan was the most-accessed feature, with almost 90% (920/1033) of the sample accessing it by the end of the study period. Other commonly accessed features included the cigarette tracker, the Beat Your Smoking Triggers, the BecomeAnEX.org community, and the separation exercises. When asked whether BecomeAnEX.org was helpful in their efforts to quit smoking, almost 90% of study participants at each of the 3 follow-up periods said yes: 1 month=86.12% (614/713), 3 month=88.12% (549/623), and 6 month=90.47% (503/556). When asked which feature (apart from the overall My Quit Plan summary page) was most helpful, respondents were most likely to endorse the cigarette tracker (20.50%, 107/522), the separation exercises (23.56%, 123/522), and the BecomeAnEX.org community (14.75%, 77/522). The features deemed least helpful by respondents at the last follow-up included the quit smoking resources (2.68%, 14/522), the nicotine addiction videos (3.64%, 19/522), and the quit smoking support exercises (4.98%, 26/522). Data indicated that use of the features on BecomeAnEX.org occurred early upon registering because there was only minimal increases (<5%) in participants accessing new features beyond the initial 1-month period. Finally, although not recorded by Omniture SiteCatalyst, survey data showed that 10% of participants reported visiting the BecomeAnEX Facebook page, and 3% visited the BecomeAnEX Twitter account over the study period.

Quit Rates

At 6 months, 7- and 30-day abstinence in the responder-only sample was 20.68% and 18.35%, respectively (see Table 2). Using ITT analysis, the rates of 7- and 30-day abstinence were 11.13% and 9.87%, respectively. Rates of abstinence increased dramatically as a function of website utilization: both 7- and 30-day abstinence rates were 10% or less for participants who had visited BecomeAnEX.org only 1 or 2 times, but jumped to 32.82% and 29.34% for those who had visited the website 3 or more times during the study period.

Table 2. Smoking behavior of survey responders over the course of the study (N=1033).

Quit rates	Baseline[a]	Follow-up		
		1 month n=713	3 months n=623	6 months n=556
Quit attempts, %				
Yes	72.22	73.38	82.48	82.73
No	27.78	26.62	17.52	17.27
Number of quit attempts, mean (SE)	3.06 (0.20)	2.12 (0.16)	3.24 (0.22)	3.52 (0.28)
Responder quit rates, %				
7-day abstinence	n/a	13.18	18.14	20.68
30-day abstinence	n/a	5.19	14.93	18.35
Intent-to-treat quit rates [b]				
7-day abstinence	n/a	9.10	10.94	11.13
30-day abstinence	n/a	3.58	9.00	9.87

[a] Smoking status at baseline is from BecomeAnEX.org registration data.

[b] Intent-to-treat treats all nonresponders as current smokers.

Characteristics of Respondents Abstinent at 6-months

Respondents abstinent for at least 7 days at the 6-month follow-up were more likely than those currently smoking to be non-Hispanic white (80.00% vs 62.36%, *P*<.001), less likely to be Hispanic (6.09% vs 23.36%, *P*<.001), less likely to have at least a college education (17.39% vs 31.29%, *P*=.003), less likely to have used the website only once (19.13% vs 41.04%, *P*<.001), and more likely to have accessed the BecomeAnEX.org community page (55.65% vs 35.37%, *P*<.001). Of respondents who accessed the BecomeAnEX.org community, those who were abstinent for at least 7 days at the 6-month follow-up, visited more often during the study period (mean 31.7, SE 12.7 vs mean 5.26, SE 2.62, *P*=.045] than those currently smoking at 6 months . There were no statistically significant differences by age or gender between those abstinent at the 6-month follow-up and those currently smoking.

Generalized Estimating Equations Models: Website Utilization and Quit Outcomes

Results of the GEE regressions examining the association between website usage and quit behavior are displayed in Table 3. Although there was no statistically significant association between visits to the website and quit attempts, visits to the website was positively associated with both 7- and 30-day abstinence over time in adjusted analyses (both *P*<.001). When the most-accessed Web features were also put into the model (Model 2), visits to the website was still significantly associated with both 7-day (*P*<.001) and 30-day abstinence (*P*=.008). Accessing the cigarette tracker was negatively associated with quit attempts over time, such that those accessing this feature more often were less likely to make a quit attempt. Those who accessed the separation exercises were 2 or more times as likely to make a quit attempt (odds ratio [OR] 2.05, 95% CI 1.00-4.19, *P*=.05) as compared to those who did not access these exercises. For the outcome of 7-day abstinence, those who accessed the BecomeAnEX.org community once had an increased odds of 7-day abstinence (OR 1.74, 95% CI 1.13-2.67. *P*=.012), as compared to those who did not access the community. This was increased even further among those who accessed the BecomeAnEX.org community 2 or more times (OR 2.22, 95% CI 1.34-3.69, *P*=.002) Accessing the separation exercises once (OR 1.58, 95% CI 1.01-2.48, *P*=.045) or at least twice (OR 1.91, 95% CI, 1.00-3.65, *P*=.049) were also significantly related to 7-day abstinence. Increases in the odds of 30-day abstinence were only associated with assessing the BecomeAnEX.org community at least 2 times (OR 2.42, 95% CI 1.35-4.34, *P*=.003). However, the association between the separation exercises and abstinence remained borderline significant.

Table 3. Association between website usage and quit behavior over time using generalized estimating equations (GEE).

Regression models	Quit attempts		7-Day abstinence		30-Day abstinence	
	OR (95% CI)	*P*	OR (95% CI)	*P*	OR (95% CI)	*P*
Model 1 [a]						
Visits to the website[b]	1.09 (0.94-1.28)	.26	2.04 (1.75-2.39)	<.001	1.73 (1.47-2.05)	<.001
Model 2 [a]						
Visits to the website[b]	1.10 (0.88-1.38)	.39	1.55 (1.26-1.91)	<.001	1.36 (1.08-1.70)	.008
Use of Community feature						
1 vs 0 times	0.79 (0.53-1.17)	.24	1.74 (1.13-2.67)	.01	1.37 (0.81-2.30)	.24
≥2 vs 0 times	1.06 (0.62-1.83)	.82	2.22 (1.34-3.69)	.002	2.42 (1.35-4.34)	.003
Use of Cigarette Tracker feature						
1 vs 0 times	0.62 (0.40-0.96)	.03	1.27 (0.48-1.21)	.33	0.97 (0.56-1.70)	.30
≥2 vs 0 times	0.42 (0.22-0.80)	.01	0.82 (0.28-1.01)	.55	0.81 (0.38-1.72)	.13
Use of Beat Triggers exercise feature						
1 vs 0 times	1.36 (0.87 2.14)	.18	1.16 (0.72-1.88)	.54	0.87 (0.50-1.52)	.63
≥2 vs 0 times	1.20 (0.60-2.37)	.61	1.20 (0.61-2.36)	.60	0.86 (0.39-1.88)	.70
Use of Separation exercise feature						
1 vs 0 times	0.99 (0.65-1.53)	.97	1.58 (1.01-2.48)	.05	1.67 (0.98-2.83)	.06
≥2 vs 0 times	2.05 (1.00-4.19)	.05	1.91(1.00-3.65)	.05	1.99 (0.94-4.22)	.07

[a] All models adjusted for demographics, nicotine dependence, baseline quit attempts, peer smoking, household smoking, motivation to quit, positive and negative social support, health status, advice to quit from a health care provider, use of at least one cessation aide, baseline depression, baseline perceived stress, having a partner who smokes, and frequency of use of the Internet. Model 1 includes all covariates plus visits to the website. Model 2 includes all covariates, visits to the website, and use of specific BecomeAnEX.org features.

[b] Represented as the log of total visits to the BecomeAnEX.org website over the study period.

Predictors of Website Utilization

The final analysis examined the predictors of website usage determined in the previous models to be associated with quit behavior, namely visits to the website and use of the BecomeAnEX.org community and separation exercises (Table 4). Data show that there were few predictors of website visits. Visits were less likely among Hispanics (beta coefficient=–0.18, SE=0.08, *P*=.03) as compared to non-Hispanic whites, and increased as function of motivation to quit (beta coefficient=0.03, SE=0.01, *P*=.02). Accessing the BecomeAnEX.org community was predicted by higher education, positive partner support at baseline (OR 1.06, 95% CI 1.01-1.10, *P*=.02), and motivation to quit (OR 1.26, 95% CI 1.16-1.37, *P*<.001). Higher reported stress at baseline was also associated with accessing the community (OR 1.05, 95% CI 1.00-1.11, *P*=.48). Negative partner support at baseline decreased the probability of accessing the community (OR 0.91, 95% CI 0.88-0.95, *P*<.001). There were many predictors of accessing the separation exercises. Separation exercises were more likely to be accessed in respondents 18 to 24 years of age, those with higher education, and those with increased motivation to quit (OR 1.37, 95% CI 1.25-1.49, *P*<.001), but less likely in non-Hispanic blacks (OR 0.37, 95% CI 0.25-0.56, *P*<.001) and those of other race/ethnicities (OR 0.42, 95% CI 0.20-0.86, *P*=.02) as compared with non-Hispanic whites and those with negative partner support at baseline (OR 0.96, 95% CI 0.93-1.00, *P*=.29). Those who had made at least 1 quit attempt (OR 0.73, 95% CI 0.54-1.00, *P*=.05) were also less likely to access the separation exercises, although this was of borderline significance.

Weighted Analyses

The 7- and 30-day weighted abstinence rates over the study period were almost identical to the unweighted rates, at 21.83% and 17.68%, respectively. Results were also similar on the measures regarding helpfulness of the website, use of Web features, and smoking behavior (see Multimedia Appendix 1). The first GEE regression model (Model 1) examining the association between website utilization and quit behavior was similar to the unweighted analyses. However, there were several differences in Model 2. Although the association between visits to the website and both 7-day abstinence (OR 1.75, 95% CI 1.35-2.28, *P*<.001) and 30-day abstinence (OR 1.54, 95% CI 1.17-2.02, *P*=.002) was similar, there was no longer a statistically significant relationship between any of the website features and quit behavior (see Multimedia Appendix 1).

Table 4. Demographic characteristics and smoking behavior with use of BecomeAnEX.org website and its associated features (unweighted).

Demographic variables	Visits to website		Use of the community		Separation exercises	
	Beta coefficient (SE)	*P*	OR (95% CI)	*P*	OR (95% CI)	*P*
Gender						
Male	–0.3 (0.05)	.56	0.88 (0.68-1.15)	.36	0.75 (0.57-0.97)	.03
Female	REF		REF		REF	
Race/ethnicity						
Non-Hispanic white	REF		REF		REF	
Non-Hispanic black	–0.11 (0.07)	.13	0.84 (0.58-1.22)	.37	0.37 (0.25-0.56)	<.001
Hispanic	–0.18 (0.08)	.03	0.72 (0.44-1.19)	.20	0.74 (0.46-1.20)	.22
Other	–0.02 (0.13)	.85	0.59 (0.28-1.22)	.16	0.42 (0.20-0.86)	.02
Age, years						
18-24	REF		REF		REF	
25-44	–0.13 (0.08)	.10	0.86 (0.56-1.32)	.48	0.56 (0.37-0.84)	.005
45-64	–0.08 (0.08)	.32	1.41 (0.90-2.22)	.14	0.51 (0.32-0.79)	.002
≥65	–0.22 (0.15)	.15	0.79 (0.34-1.84)	.58	0.35 (0.15-0.84)	.02
Education						
Less than high school	–0.04 (0.10)	.69	1.47 (0.82-2.62)	.20	0.32 (0.15-0.68)	.003
High school diploma/GED	REF		REF		REF	
Some College	0.07 (0.06)	.20	1.51 (1.09-2.10)	.01	1.49 (1.08-2.06)	.02
College degree or more	0.10 (0.07)	.17	1.59 (1.07-2.36)	.02	1.80 (1.22-2.65)	.003
Baseline smoking variables						
Past-year quit attempts	–0.09 (0.06)	.10	0.82 (0.60-1.10)	.18	0.73 (0.54-1.00)	.05
Fagerstrom Test for Nicotine Dependence	0.02 (0.02)	.46	1.05 (0.94-1.18)	.36	0.90 (0.80-1.00)	.06
Motivation to quit (1-10)	0.03 (0.01)	.02	1.26 (1.16-1.37)	<.001	1.37 (1.25-1.49)	<.001
Use of at least one cessation aide	0.01 (0.05)	.90	0.89 (0.68-1.17)	.40	1.16 (0.88-1.52)	.29
Health status						
Excellent	0.01 (0.14)	.93	1.16 (0.51-2.64)	.73	0.46 (0.20-1.04)	.06
Very good	0.06 (0.11)	.57	1.93 (1.02-3.66)	.04	0.90 (0.49-1.68)	.74
Good	0.08 (0.11)	.48	1.62 (0.88-2.97)	.12	0.91 (0.51-1.65)	.77
Fair	0.07 (0.11)	.56	1.54 (0.83-2.88)	.98	0.80 (0.43-1.47)	.47
Poor	REF		REF		REF	
Advised to quit from a health care professional	0.04 (0.06)	.51	1.33 (0.91-1.93)	.14	0.90 (0.63-1.29)	.58
Baseline sadness	–0.02 (0.02)	.34	1.00 (0.88-1.13)	.98	0.94 (0.83-1.07)	.34
Baseline perceived stress	0.01 (0.01)	.25	1.05 (1.00-1.11)	.05	1.01 (0.96-1.07)	.69
Baseline partner support (PIQ)						
Positive support	0.002 (0.01)	.83	1.06 (1.01-1.10)	.02	1.02 (0.97-1.06)	.48
Negative support	–0.008 (0.01)	.22	0.91 (0.88-0.95)	<.001	0.96 (0.93-1.00)	.03
Peer smoking	-0.01 (0.01)	.59	1.03 (0.95-1.11)	.55	1.04 (0.96-1.13)	.32
Has a partner who smokes	–0.08 (0.05)	.89	0.85 (0.63-1.14)	.27	0.95 (0.71-1.28)	.76
Household smoking	–0.04 (0.05)	.45	1.04 (0.79-1.37)	.80	0.87 (0.66-1.14)	.32
Frequency of use of the Internet[a]	–0.03 (0.03)	.33	0.92 (0.76-1.12)	.42	1.16 (0.96-1.40)	.12

[a] Modeled as a continuous variable, the higher the number, the less often they use.

Discussion

This observational prospective study adds to the growing body of literature examining the effectiveness of an interactive Web-based smoking cessation intervention in promoting quit behavior and presents new information on factors predicting general utilization and use of specific features of this website. Abstinence rates were approximately 20.68% for 7-day and 18.35% for 30-day abstinence at 6 months in both weighted and unweighted analyses, and 11.13% and 9.87% for 7- and 30-day abstinence in the ITT analysis. These rates were similar to quit rates achieved in other Web-based smoking cessation interventions [18] and highlight the importance of finding ways to increase repeat visits, engagement, and retention of new users, both in the general population and those participants involved in a real-world evaluation study of a Web-based intervention.

Adjusting for a host of demographic factors, smoking history, and psychosocial variables thought to potentially influence quit behavior, the number of visits to BecomeAnEX.org significantly predicted both 7- and 30-day abstinence. Even after accounting for the independent and statistically significant effects of accessing the BecomeAnEX.org community and separation exercises, number of visits still predicted abstinence. Motivation to quit was a key predictor of visiting the website and accessing both the community and the separation exercises, whereas negative partner support decreased the likelihood of accessing both of these features.

Results presented here are in accordance with several studies that have reported significant associations between number of visits to a Web-based smoking cessation intervention and abstinence [28-32] and build on existing evidence suggesting the importance of interactive features in promoting cessation [18,53] as well as other behavioral outcomes.[54-56] Among the top features accessed on the website, the BecomeAnEX.org community most significantly predicted abstinence in the sample, followed by the separation exercises. The remaining interactive features—the Cigarette Tracker, Beat Your Triggers exercises, and the Set a Quit Date feature—did not promote abstinence. In fact, accessing the cigarette tracker was negatively associated with quit attempts, potentially because participants accessing this feature were still contemplating and/or preparing for a behavior change and not ready to quit. It may work, therefore, to move someone down the pathway toward behavior change and/or it may be that other elements are needed in this feature to prompt a smoker to quit. These findings suggest that interactivity itself may not be sufficient to change behavior. The challenge is to identify how features such as the community and separation exercises work to promote abstinence, to consider whether such features successfully help move smokers at all stages of behavior change toward making a quit attempt, and then design Web-based interventions that capitalize on this information.

Data presented here also suggest that the results achieved in a study sample may differ from the broader population of registrants.Although most of the weighted results were similar to the unweighted results, associations between BecomeAnEX.org community utilization and abstinence and between utilization of the separation exercises and abstinence were in the same direction, but the effects were not statistically significant in weighted analyses. This may be because of underlying differences in motivation to quit between general registrants and a sample that must go through multiple steps of recruitment before entering into a study. It could also reflect differences in social support between registrants and the sample that influenced the use of the interactive features and potentially could influence how these features promoted cessation.

Maximizing the effectiveness of any cessation intervention is dependent on reaching the target audience, but also engaging and motivating the smoker to take concrete steps toward quitting. Data from this study and others suggest that engagement may be the key driver of success, but the nature of engagement and the mechanisms underlying the effectiveness of engagement in promoting abstinence may differ across subgroups. For example, data presented here show that those abstinent at 6 months were more likely to be white, less likely to be Hispanic, and less likely to be at least college-educated. The reasons for this are unclear, but may be because of the layout, design, language, and features embedded in the quit program. Further research is needed to clarify why certain populations appear to benefit more than others from this and other Web-based smoking cessation interventions. Because one intervention will never appropriately cater to the needs of all subpopulations, it is essential for nonprofit, for-profit, and governmental groups to collaboratively develop and promote complementary and synergistic programs to maximize effectiveness across the population as a whole.

There are several limitations to this study. First, results are limited to those respondents who completed at least 1 follow-up survey. Response rates on the follow-up surveys (69.41%, 60.31%, and 53.92% at 1, 3, and 6 months, respectively) were as good or better than those recorded for many Web-based studies [20,26,27,29,32,57] (eg, 39.4% online and 38.2% by telephone [58] and 48% and 45% at 3 and 6 months [59]). Although GEE was used to maximize the information obtained from the prospective data, these data support the need to employ techniques to minimize attrition to provide the most valid estimates of the intervention's effectiveness [60-62]. Second, although software tracked respondents' activity online, data presented here are limited to what features the respondent accessed and do not necessarily indicate full utilization of a specific feature of the site. For example, they could have accessed a page on the BecomeAnEX.org community but not posted or responded to a thread. Follow-up studies will need to address this further by using more in-depth tracking metrics. Third, individual-level utilization of the mobile BecomeAnEX.org site could not be characterized by using the Omniture SiteCatalyst software. Since the mobile site was not heavily promoted or used during the trial, however, it is unlikely to have significantly biased results. Fourth, it is possible that the weighting scheme did not entirely correct for differences in website utilization, motivation to quit, or smoking behaviors between the study sample and the larger BecomeAnEX.org population. However, the demographic variables used for weighting were shown to be significantly correlated with motivation to quit, quit attempts, cigarettes per day, and website utilization, thereby helping to correct any differences that exist between these populations. Fifth, the study period was only 6 months and outcomes may have changed if respondents were followed for a longer period of time. Six-month follow-up for this type of observational study is standard in the literature [18,37]. Sixth, there was no control group used in this study, so the overall effectiveness versus placebo could not be determined. Finally, differences between the study sample and the larger population of BecomeAnEX.org members may limit generalizability of the results. These challenges to generalizability confront any real-world study that focuses on practical trials from an implementation-dissemination research methods perspective [38,40].

Tobacco use causes nearly 6 million deaths and costs hundreds of billions of dollars worldwide each year [63]. Interventions are needed that are low cost, accessible, sustainable, and can reach large populations [64]. This study adds to the growing literature that seeks to understand how to maximize the effectiveness of Web-based smoking cessation interventions. It is the first observational study that weights the data to obtain representative estimates of the larger population of website registrants, thereby addressing a key question of the effectiveness of Web-based smoking cessation interventions in the greater population of users. This study also employs website tracking software to generate reliable estimates of website utilization and also controls for many potential confounders that are not often considered in such analyses, including stress, depression, and both positive and negative partner support. Results suggest that the primary driver of success in Web-based smoking cessation interventions is engagement through multiple visits, and involvement in the community and other interactive features. Findings emphasize the importance of increasing the interaction of new users with other smokers wanting to quit and finding ways to increase repeat visits and retention. This study also addresses a key scientific question of the type of smoker most likely to use and benefit from Web-based smoking cessation interventions. Data suggest that public health organizations looking to affect the greatest change in behavior may achieve maximal results by marketing these resources to smokers who are highly motivated to quit. Future research should further examine the variation in effectiveness across demographic subgroups, how use of interactive features promotes cessation, mechanisms underlying effectiveness of these features, whether there are synergies among different components of the website, and how results gained through evaluations of Web-based cessation interventions apply to the greater population of registered users. This information can help advance the development and use of Web-based smoking cessation interventions for the greatest impact on population-level smoking rates.

Conflicts of Interest

All authors work for Legacy, the company that funded the development, design, and implementation of the Web-based smoking cessation site evaluated in this manuscript. Dr. Cobb is a consultant to MeYouHealth whose parent company's product line includes an online tobacco cessation intervention.

Multimedia Appendix 1

Weighted analysis tables.

[PDF File (Adobe PDF File), 114KB - jmir_v15i1e14_app1.pdf]

References

1. Centers for Disease Control and Prevention (CDC). Vital signs: current cigarette smoking among adults aged ≥18 years--United States, 2005-2010. MMWR Morb Mortal Wkly Rep 2011 Sep 9;60(35):1207-1212 [FREE Full text] [Medline: 21900875]
2. Centers for Disease Control and Prevention (CDC). Quitting smoking among adults--United States, 2001-2010. MMWR Morb Mortal Wkly Rep 2011 Nov 11;60(44):1513-1519 [FREE Full text] [Medline: 22071589]
3. Internet World Stats. Internet usage statistics: The Internet big picture 2012 URL: http://www.internetworldstats.com/stats.htm [accessed 2012-06-29] [WebCite Cache ID 6AVBK47sh]
4. Fox S. Online health search 2006. Washington, DC: Pew Internet & American Life Project; 2006 Oct 29. URL: http://pewinternet.org/Reports/2006/Online-Health-Search-2006.aspx [accessed 2012-06-16] [WebCite Cache ID 6DiUCs73p]
5. Cobb NK. Online consumer search strategies for smoking-cessation information. Am J Prev Med 2010 Mar;38(3 Suppl):S429-S432. [doi: 10.1016/j.amepre.2009.12.001] [Medline: 20176318]
6. Abrams DB, Orleans CT, Niaura RS, Goldstein MG, Prochaska JO, Velicer W. Integrating individual and public health perspectives for treatment of tobacco dependence under managed health care: a combined stepped-care and matching model. Ann Behav Med 1996;18(4):290-304. [doi: 10.1007/BF02895291] [Medline: 18425675]
7. Abrams D. Comprehensive smoking cessation policy for all smokers: systems integration to save livesmoney. In: Ending the Tobacco Problem: A Blueprint for the Nation. Washington, DC: National Academies Press; 2007.
8. Fiore MC, Jaén CR, Baker TB, Bailey WC, Benowitz NL, Curry SJ, et al. Treating Tobacco Use and Dependence: 2008 Update (Clinical Practice Guideline). Rockville, MD: Department of Health and Human Services Public Health Services; 2008. URL: http://www.ahrq.gov/clinic/tobacco/treating_tobacco_use08.pdf [accessed 2013-01-17] [WebCite Cache ID 6DkJtR2ZO]

9. Whittaker R, Maddison R, McRobbie H, Bullen C, Denny S, Dorey E, et al. A multimedia mobile phone-based youth smoking cessation intervention: findings from content development and piloting studies. J Med Internet Res 2008;10(5):e49 [FREE Full text] [doi: 10.2196/jmir.1007] [Medline: 19033148]
10. Whittaker R, Dorey E, Bramley D, Bullen C, Denny S, Elley CR, et al. A theory-based video messaging mobile phone intervention for smoking cessation: randomized controlled trial. J Med Internet Res 2011;13(1):e10 [FREE Full text] [doi: 10.2196/jmir.1553] [Medline: 21371991]
11. Free C, Knight R, Robertson S, Whittaker R, Edwards P, Zhou W, et al. Smoking cessation support delivered via mobile phone text messaging (txt2stop): a single-blind, randomised trial. Lancet 2011 Jul 2;378(9785):49-55 [FREE Full text] [doi: 10.1016/S0140-6736(11)60701-0] [Medline: 21722952]
12. Kramish Campbell M, Meier A, Carr C, Enga Z, James AS, Reedy J, et al. Health behavior changes after colon cancer: a comparison of findings from face-to-face and on-line focus groups. Fam Community Health 2001 Oct;24(3):88-103. [Medline: 11563947]
13. Stoddard JL, Augustson EM, Moser RP. Effect of adding a virtual community (bulletin board) to smokefree.gov: randomized controlled trial. J Med Internet Res 2008;10(5):e53 [FREE Full text] [doi: 10.2196/jmir.1124] [Medline: 19097974]
14. Brendryen H, Kraft P. Happy ending: a randomized controlled trial of a digital multi-media smoking cessation intervention. Addiction 2008 Mar;103(3):478-84; discussion 485. [doi: 10.1111/j.1360-0443.2007.02119.x] [Medline: 18269367]
15. Cobb NK, Graham AL, Abrams DB. Social network structure of a large online community for smoking cessation. Am J Public Health 2010 Jul;100(7):1282-1289 [FREE Full text] [doi: 10.2105/AJPH.2009.165449] [Medline: 20466971]
16. Myung SK, McDonnell DD, Kazinets G, Seo HG, Moskowitz JM. Effects of Web- and computer-based smoking cessation programs: meta-analysis of randomized controlled trials. Arch Intern Med 2009 May 25;169(10):929-937. [doi: 10.1001/archinternmed.2009.109] [Medline: 19468084]
17. Civljak M, Sheikh A, Stead LF, Car J. Internet-based interventions for smoking cessation. Cochrane Database Syst Rev 2010(9):CD007078. [doi: 10.1002/14651858.CD007078.pub3] [Medline: 20824856]
18. Shahab L, McEwen A. Online support for smoking cessation: a systematic review of the literature. Addiction 2009 Nov;104(11):1792-1804. [doi: 10.1111/j.1360-0443.2009.02710.x] [Medline: 19832783]
19. Muñoz RF, Barrera AZ, Delucchi K, Penilla C, Torres LD, Pérez-Stable EJ. International Spanish/English Internet smoking cessation trial yields 20% abstinence rates at 1 year. Nicotine Tob Res 2009 Sep;11(9):1025-1034 [FREE Full text] [doi: 10.1093/ntr/ntp090] [Medline: 19640833]
20. Swartz LH, Noell JW, Schroeder SW, Ary DV. A randomised control study of a fully automated internet based smoking cessation programme. Tob Control 2006 Feb;15(1):7-12 [FREE Full text] [doi: 10.1136/tc.2003.006189] [Medline: 16436397]
21. 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]
22. Strecher VJ, Shiffman S, West R. Randomized controlled trial of a web-based computer-tailored smoking cessation program as a supplement to nicotine patch therapy. Addiction 2005 May;100(5):682-688. [doi: 10.1111/j.1360-0443.2005.01093.x] [Medline: 15847626]
23. Strecher VJ, McClure JB, Alexander GL, Chakraborty B, Nair VN, Konkel JM, et al. Web-based smoking-cessation programs: results of a randomized trial. Am J Prev Med 2008 May;34(5):373-381 [FREE Full text] [doi: 10.1016/j.amepre.2007.12.024] [Medline: 18407003]
24. An LC, Klatt C, Perry CL, Lein EB, Hennrikus DJ, Pallonen UE, et al. The RealU online cessation intervention for college smokers: a randomized controlled trial. Prev Med 2008 Aug;47(2):194-199. [doi: 10.1016/j.ypmed.2008.04.011] [Medline: 18565577]
25. Elfeddali I, Bolman C, Candel MJ, Wiers RW, de Vries H. Preventing smoking relapse via Web-based computer-tailored feedback: a randomized controlled trial. J Med Internet Res 2012;14(4):e109 [FREE Full text] [doi: 10.2196/jmir.2057] [Medline: 22903145]
26. Mason D, Gilbert H, Sutton S. Effectiveness of web-based tailored smoking cessation advice reports (iQuit): a randomized trial. Addiction 2012 Dec;107(12):2183-2190. [doi: 10.1111/j.1360-0443.2012.03972.x] [Medline: 22690882]
27. Smit ES, de Vries H, Hoving C. Effectiveness of a Web-based multiple tailored smoking cessation program: a randomized controlled trial among Dutch adult smokers. J Med Internet Res 2012;14(3):e82 [FREE Full text] [doi: 10.2196/jmir.1812] [Medline: 22687887]
28. Strecher VJ, McClure J, Alexander G, Chakraborty B, Nair V, Konkel J, et al. The role of engagement in a tailored web-based smoking cessation program: randomized controlled trial. J Med Internet Res 2008;10(5):e36 [FREE Full text] [doi: 10.2196/jmir.1002] [Medline: 18984557]
29. Pike KJ, Rabius V, McAlister A, Geiger A. American Cancer Society's QuitLink: randomized trial of Internet assistance. Nicotine Tob Res 2007 Mar;9(3):415-420. [doi: 10.1080/14622200701188877] [Medline: 17365773]
30. Rabius V, Pike KJ, Wiatrek D, McAlister AL. Comparing internet assistance for smoking cessation: 13-month follow-up of a six-arm randomized controlled trial. J Med Internet Res 2008;10(5):e45 [FREE Full text] [doi: 10.2196/jmir.1008] [Medline: 19033146]

31. Japuntich SJ, Zehner ME, Smith SS, Jorenby DE, Valdez JA, Fiore MC, et al. Smoking cessation via the internet: a randomized clinical trial of an internet intervention as adjuvant treatment in a smoking cessation intervention. Nicotine Tob Res 2006 Dec;8 Suppl 1:S59-S67. [Medline: 17491172]
32. Stoddard JL, Augustson EM, Moser RP. Effect of adding a virtual community (bulletin board) to smokefree.gov: randomized controlled trial. J Med Internet Res 2008;10(5):e53 [FREE Full text] [doi: 10.2196/jmir.1124] [Medline: 19097974]
33. Cobb NK, Graham AL, Bock BC, Papandonatos G, Abrams DB. Initial evaluation of a real-world Internet smoking cessation system. Nicotine Tob Res 2005 Apr;7(2):207-216 [FREE Full text] [doi: 10.1080/14622200500055319] [Medline: 16036277]
34. Graham AL, Cobb NK, Raymond L, Sill S, Young J. Effectiveness of an internet-based worksite smoking cessation intervention at 12 months. J Occup Environ Med 2007 Aug;49(8):821-828. [doi: 10.1097/JOM.0b013e3180d09e6f] [Medline: 17693778]
35. Seidman DF, Westmaas JL, Goldband S, Rabius V, Katkin ES, Pike KJ, et al. Randomized controlled trial of an interactive internet smoking cessation program with long-term follow-up. Ann Behav Med 2010 Feb;39(1):48-60. [doi: 10.1007/s12160-010-9167-7] [Medline: 20177844]
36. Zbikowski SM, Hapgood J, Smucker Barnwell S, McAfee T. Phone and web-based tobacco cessation treatment: real-world utilization patterns and outcomes for 11,000 tobacco users. J Med Internet Res 2008;10(5):e41 [FREE Full text] [doi: 10.2196/jmir.999] [Medline: 19017583]
37. Hutton HE, Wilson LM, Apelberg BJ, Tang EA, Odelola O, Bass EB, et al. A systematic review of randomized controlled trials: Web-based interventions for smoking cessation among adolescents, college students, and adults. Nicotine Tob Res 2011 Apr;13(4):227-238. [doi: 10.1093/ntr/ntq252] [Medline: 21350042]
38. Glasgow RE, Green LW, Klesges LM, Abrams DB, Fisher EB, Goldstein MG, et al. External validity: we need to do more. Ann Behav Med 2006 Apr;31(2):105-108. [doi: 10.1207/s15324796abm3102_1] [Medline: 16542124]
39. Mabry PL, Olster DH, Morgan GD, Abrams DB. Interdisciplinarity and systems science to improve population health: a view from the NIH Office of Behavioral and Social Sciences Research. Am J Prev Med 2008 Aug;35(2 Suppl):S211-S224 [FREE Full text] [doi: 10.1016/j.amepre.2008.05.018] [Medline: 18619402]
40. Glasgow RE, Emmons KM. How can we increase translation of research into practice? Types of evidence needed. Annu Rev Public Health 2007;28:413-433. [doi: 10.1146/annurev.publhealth.28.021406.144145] [Medline: 17150029]
41. Levy DT, Graham AL, Mabry PL, Abrams DB, Orleans CT. Modeling the impact of smoking-cessation treatment policies on quit rates. Am J Prev Med 2010 Mar;38(3 Suppl):S364-S372 [FREE Full text] [doi: 10.1016/j.amepre.2009.11.016] [Medline: 20176309]
42. BecomeAnEx.org. About EX URL: http://www.becomeanex.org/about-ex.php [accessed 2013-01-17] [WebCite Cache ID 6DkLdvCWU]
43. Adobe. SiteCatalyst Product Overview URL: http://www.omniture.com/en/products/analytics/sitecatalyst [accessed 2012-09-07] [WebCite Cache ID 6AVAR9CxH]
44. Strecher VJ, McClure J, Alexander G, Chakraborty B, Nair V, Konkel J, et al. The role of engagement in a tailored web-based smoking cessation program: randomized controlled trial. J Med Internet Res 2008;10(5):e36 [FREE Full text] [doi: 10.2196/jmir.1002] [Medline: 18984557]
45. SRNT Subcommittee on Biochemical Verification. Biochemical verification of tobacco use and cessation. Nicotine Tob Res 2002 May;4(2):149-159. [doi: 10.1080/14622200210123581] [Medline: 12028847]
46. 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]
47. Graham AL, Papandonatos GD, Bock BC, Cobb NK, Baskin-Sommers A, Niaura R, et al. Internet- vs. telephone-administered questionnaires in a randomized trial of smoking cessation. Nicotine Tob Res 2006 Dec;8 Suppl 1:S49-S57 [FREE Full text] [Medline: 17491171]
48. Ware J, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996 Mar;34(3):220-233. [Medline: 8628042]
49. Cohen S, Williamson G. Perceived stress in a probability sample of the United States. In: Spacapam S , Oskamp S , editors. The Social Psychology of Health: Claremont Symposium on Applied Social Psychology. Newbury Park, CA: Sage; 1988.
50. Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. Biometrika 1986 Apr;73(1):13-22.
51. Hanley JA, Negassa A, Edwardes MD, Forrester JE. Statistical analysis of correlated data using generalized estimating equations: an orientation. Am J Epidemiol 2003 Feb 15;157(4):364-375 [FREE Full text] [Medline: 12578807]
52. United States Census Bureau. 2011. USA QuickFacts URL: http://quickfacts.census.gov/qfd/states/00000.html [accessed 2012-09-07] [WebCite Cache ID 6AVAjepeA]
53. Schwarzer R, Satow L. Online intervention engagement predicts smoking cessation. Prev Med 2012 Sep;55(3):233-236. [doi: 10.1016/j.ypmed.2012.07.006] [Medline: 22813919]
54. Poirier J, Cobb NK. Social influence as a driver of engagement in a web-based health intervention. J Med Internet Res 2012;14(1):e36 [FREE Full text] [doi: 10.2196/jmir.1957] [Medline: 22356829]
55. 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]

56. Couper MP, Alexander GL, Zhang N, Little RJ, 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]
57. Smit ES, Hoving C, Cox VC, de Vries H. Influence of recruitment strategy on the reach and effect of a web-based multiple tailored smoking cessation intervention among Dutch adult smokers. Health Educ Res 2012 Apr;27(2):191-199. [doi: 10.1093/her/cyr099] [Medline: 22058288]
58. Saul JE, Schillo BA, Evered S, Luxenberg MG, Kavanaugh A, Cobb N, et al. Impact of a statewide Internet-based tobacco cessation intervention. J Med Internet Res 2007;9(3):e28 [FREE Full text] [doi: 10.2196/jmir.9.4.e28] [Medline: 17942390]
59. Severson HH, Gordon JS, Danaher BG, Akers L. ChewFree.com: evaluation of a Web-based cessation program for smokeless tobacco users. Nicotine Tob Res 2008 Feb;10(2):381-391. [doi: 10.1080/14622200701824984] [Medline: 18236303]
60. 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]
61. Postel MG, de Haan HA, ter Huurne ED, van der Palen J, Becker ES, de Jong CA. Attrition in web-based treatment for problem drinkers. J Med Internet Res 2011;13(4):e117 [FREE Full text] [doi: 10.2196/jmir.1811] [Medline: 22201703]
62. Cugelman B, Thelwall M, Dawes P. Online interventions for social marketing health behavior change campaigns: a meta-analysis of psychological architectures and adherence factors. J Med Internet Res 2011;13(1):e17 [FREE Full text] [doi: 10.2196/jmir.1367] [Medline: 21320854]
63. World Health Organization. WHO Report on the Global Tobacco Epidemic 2011: Warning About the Dangers of Tobacco (Documents for Sale). 2011. URL: http://whqlibdoc.who.int/publications/2011/9789240687813_eng.pdf [accessed 2013-01-17] [WebCite Cache ID 6DkKrLPoT]
64. 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]

Abbreviations

GED: Generalized Education Development Test
GEE: generalized estimating equations
IQR: interquartile range
ITT: intent-to-treat
IRB: Institutional Review Board
OR: odds ratio
PIQ: Partner Interaction Questionnaire
SF-12: 12-Item Short Form Health Survey

Edited by G Eysenbach; submitted 23.07.12; peer-reviewed by S Zbikowski, J McClure, J Whiteley; comments to author 25.08.12; revised version received 07.09.12; accepted 25.09.12; published 28.01.13.

Please cite as:
Richardson A, Graham AL, Cobb N, Xiao H, Mushro A, Abrams D, Vallone D
Engagement Promotes Abstinence in a Web-based Cessation Intervention: Cohort Study
J Med Internet Res 2013;15(1):e14
URL: http://www.jmir.org/2013/1/e14/
doi:10.2196/jmir.2277
PMID:23353649

Original Paper

Type 1 Diabetes eHealth Psychoeducation: Youth Recruitment, Participation, and Satisfaction

Robin Whittemore[1], PhD, RN, FAAN; Sarah S Jaser[2], PhD; Melissa S Faulkner[3], DSN, RN, FAAN; Kathryn Murphy[4], PhD, RN; Alan Delamater[5], PhD; Margaret Grey[1], DrPH, RN, FAAN; TEENCOPE Research Group[1]

[1]Yale, School of Nursing, New Haven, CT, United States
[2]Vanderbilt University, Department of Pediatrics, Nashville, TN, United States
[3]University of Arizona, Tucson, AZ, United States
[4]The Children's Hospital of Philadelphia, Diabetes Center for Children, Philadelphia, PA, United States
[5]University of Miami, Miami, FL, United States

Corresponding Author:
Robin Whittemore, PhD, RN, FAAN
Yale
School of Nursing
100 Church Street South
P.O. Box 9740
New Haven, CT, 06536-0740
United States
Phone: 1 203 737 2351
Fax: 1 203 737 4480
Email: robin.whittemore@yale.edu

Abstract

Background: The Internet and other eHealth technologies offer a platform for improving the dissemination and accessibility of psychoeducational programs for youth with chronic illness. However, little is known about the recruitment process and yield of diverse samples in Internet research with youth who have a chronic illness.

Objective: The purpose of this study was to compare the demographic and clinical characteristics of youth with Type 1 diabetes on recruitment, participation, and satisfaction with 2 eHealth psychoeducational programs.

Methods: Youth with Type 1 diabetes from 4 sites in the United States were invited to participate (N=510) with 320 eligible youth consenting (mean age=12.3, SD 1.1; 55.3% female; 65.2% white; and mean A1C=8.3, SD 1.5). Data for this secondary analysis included demographic information (age, race/ethnicity, and income), depressive symptoms, and recruitment rates, including those who refused at point of contact (22.0%), passive refusers who consented but did not participate (15.3%), and those who enrolled (62.7%). Participation (80% lessons completed) and a satisfaction survey (ie, how helpful, enjoyable) were also analyzed. Chi-square or analysis of variance (ANOVA) analyses were used.

Results: There were significant differences in recruitment rates by income and race/ethnicity such that black, Hispanic, or mixed race/ethnicity and low-income youth were more likely to refuse passively compared to white and higher-income youth who were more likely to enroll (*P*<.001). Participation in program sessions was high, with 78.1% of youth completing at least 4 of 5 sessions. There were no significant differences in participation by program, age, gender, or race/ethnicity. Low-income youth were less likely to participate (*P*=.002). Satisfaction in both programs was also high (3.9 of 5). There were significant gender, race/ethnicity, and income differences, in that girls (*P*=.001), black, Hispanic, or mixed race/ethnicity youth (*P*=.02), and low-income youth (*P*=.02) reported higher satisfaction. There were no differences in satisfaction by program or age.

Conclusions: Results indicate that black, Hispanic, or mixed race/ethnicity youth and low-income youth with Type 1 diabetes are less likely to enroll in Internet-based research than white and higher-income youth; thus, creative recruitment approaches are needed. Low-income youth were less likely to participate, possibly due to access. However, once enrolled, youth of diverse race/ethnicity and low-income youth with Type 1 diabetes were as highly satisfied with the eHealth programs as white youth and those with higher income. Results suggest that eHealth programs have the potential to reach diverse youth and be appealing to them.

(J Med Internet Res 2013;15(1):e15) doi:10.2196/jmir.2170

KEYWORDS

Internet; patient participation rates; patient satisfaction; research subject recruitment youth

Introduction

Type 1 diabetes is a common chronic illness in adolescents, affecting 1 in 400 youths [1]. The racial/ethnicity distribution of Type 1 diabetes affects primarily white youth in the United States. In a series of studies, the SEARCH for Diabetes in Youth research group reported that the prevalence of Type 1 diabetes in youths is approximately 70% white, 22% Hispanic, and 8% black [2-4]. Management of Type 1 diabetes is complex, requiring frequent monitoring of blood glucose levels, symptoms, and carbohydrate intake. Daily insulin treatment (3-4 injections/day or infusion from a pump) as well as meal-to-meal adjustment of insulin dose depending on diet and activity patterns is required [5]. As youth transition to adolescence and take on greater responsibility for their Type 1 diabetes management and decision making, adherence to diabetes tasks often deteriorates [6], resulting in family conflict, psychological distress, and poor metabolic control [7,8].

Psychoeducational programs for youths and family-based programs have shown to be effective in improving psychosocial and diabetes-related outcomes [9-11]. Psychoeducational programs provide education, behavioral skills, and psychosocial support for young people and their families to learn how to optimally manage a chronic illness. However, disseminating and translating research-based programs into clinical care has been challenging because of provider and family time constraints, as well as cost [12].

The Internet and other eHealth technologies offer a platform for improving the dissemination and accessibility of psychoeducational programs for youth with Type 1 diabetes. Access to the Internet is increasingly available nationwide, with 94% of youth online regularly [13]. Approximately 90% of young people of all demographic and socioeconomic categories have access to the Internet [13]. Thus, eHealth interventions have the potential to reach a diverse group of youths. Programs provided on the Internet can include psychoeducational content, interactive learning, immediate feedback, and social networking [14,15].

Psychoeducational programs delivered via computer-based Internet access have demonstrated efficacy in youths with various chronic illnesses, leading to improved knowledge, symptoms, health outcomes, and quality of life [14,15]. With respect to youths with Type 1 diabetes, an eHealth self-management program with a focus on problem solving and social networking demonstrated improved self-management and problem solving in youth who completed the program compared to a control group [16]. An Internet coping skills training program, developed by our research team, did not demonstrate differential improvements in metabolic control and diabetes-related outcomes compared to an Internet diabetes education program, but youths in both groups reported significantly increased self-care autonomy, higher diabetes self-efficacy, and improved overall quality of life over time [17]. The Internet, therefore, represents a potentially efficient and effective delivery platform for psychoeducational programs for youths with Type 1 diabetes and other chronic illnesses. When evaluated, high satisfaction with eHealth programs have also been reported [18].

Despite the numerous benefits of eHealth programs for youths with chronic illnesses, concerns have been raised about the “digital divide” and Internet access for youths of diverse race and ethnicity and those living in low-income families. Although the vast majority of youths are online, access is higher in white youths and those who live in high-income families [13]. White youths and youths in high-income families are more likely to have online access at home (96%) and go online more frequently compared with black youths (92%), Hispanic youths (87%), and youths from low-income families (86%) [13]. A positive relationship between socioeconomic status and computer-based Internet use has been demonstrated in diverse middle school students [19] and a diverse pediatric clinic population [20].

Challenges of recruiting youths of diverse race and ethnicities for research are well established [21,22]. Issues about the perceived value of research, access to research for families of all strata of society, and cumbersome informed consent procedures have been documented [22]. However, little is known about the recruitment process and yield of diverse samples in Internet research with youths who have a chronic illness.

In addition to concerns regarding access to eHealth research, participation in eHealth programs has varied considerably across studies [14]. For example, in an eHealth program for youth with asthma, participants did not complete self-monitoring on 60% of study days [23]. In an eHealth program for depression, only 30% of youths completed 50% or more of the program modules [24]. In the eHealth problem-solving program for youth with Type 1 diabetes, the mean number of modules completed was 5.22 (of 8), with only 63% of youths completing all modules [18]. Participation in eHealth programs for youths typically decreases over the course of the study [23,25] and higher participation has been associated with more positive outcomes. For example, school-aged youths who had greater participation in an eHealth obesity prevention program demonstrated improved outcomes compared to youths with less participation [26]. A structured environment (ie, school vs home) may improve youth participation in eHealth programs. Youths who participated in a school-based eHealth program for depression had almost a 10-fold higher completion rate for modules and program exercises compared to youths who participated in the same program delivered as open access online [27].

Factors associated with eHealth program participation have begun to be identified. Girls have demonstrated greater participation compared to boys [18,27]. Increased depressive symptoms may also influence participation, although this effect may vary depending on the characteristics of the program, severity of symptoms, and the type of chronic illness. For example, in one study evaluating an eHealth program for depression treatment, less participation was reported in youths with higher depressive symptoms at baseline [24]. In contrast, in another study evaluating an eHealth depression prevention program, the authors reported that youth with higher depressive symptoms had greater participation in the program [28].

In summary, recruitment, participation, and satisfaction with eHealth programs have the potential to influence eHealth program outcomes and generalizability of results. Yet, little research has been undertaken to systematically evaluate the recruitment, participation, and satisfaction of eHealth programs. Therefore, the purpose of this study was to compare the demographic and clinical characteristics of youth with Type 1 diabetes on recruitment, participation, and satisfaction with 2 eHealth psychoeducational programs. Specifically, recruitment, participation, and satisfaction were compared by age, gender, race/ethnicity, household income, metabolic control, and depressive symptoms.

Method

The current study is a secondary analysis of data from a clinical trial evaluating the effect of an Internet coping skills training program (TEENCOPE) compared to an Internet diabetes health education program (Managing Diabetes) for youth with Type 1 diabetes. Each program consisted of 5 sessions with content tailored to adolescents with Type 1 diabetes. TEENCOPE used a cast of ethnically diverse characters (youth with Type 1 diabetes) and a graphic novel format to model common problematic social situations (ie, parent conflict) and different coping skills to solve the problems (Figures 1 and 2). Managing Diabetes used visuals and a highly interactive interface that allowed adolescents to actively problem-solve diabetes self-management situations (Figure 3 and 4) [29].

A convenience sample was recruited from 4 university-affiliated clinical sites that included Children's Hospital of Pennsylvania, Philadelphia, PA; University of Arizona, Tucson, AZ; University of Miami, Miami, FL; and Yale University, New Haven, CT. Inclusion criteria were: youth diagnosed with Type 1 diabetes for at least 6 months, aged 11 to 14 years, with no other significant medical problem, school grade appropriate to age within 1 year, able to speak and write English, and access to high-speed Internet at home, school, community, or clinic.

Figure 1. TEENCOPE screenshot displaying home page image, which is updated as teen progresses through the program.

"Last week you learned about stress and stress management. I downloaded the "Imagining a Peaceful Place" to my IPOD and practiced it before I went to sleep. It really helped! Hope that you found a new relaxation exercise that you like to do. I can't believe this is already the last session. Today we'll be talking about Conflict Resolution and saying good-bye to your group."

Last session—Conflict Resolution. See how the TeenCope crew resolve conflict and learn your conflict style!

Sponsored by NIH: R01NR4009
All contents © Copyright 2008, Yale University School of Nursing. All rights reserved.
Problems or questions: teencope@yale.edu

Figure 2. TEENCOPE screenshot displaying graphic novel format for conflict resolution segment.

Session 5 » Conflict Resolution

The Nature of Conflict

When you think about **conflict**, what comes to mind?

When I think about conflict, I think about disagreements and arguing with my mother.

War is when you have conflict on a really big scale.

Sometimes conflict ends up in an ugly fight that badly affects friendships.

Sometimes conflict is necessary. It helps clear the air.

Sponsored by NIH: R01NR4009
All contents © Copyright 2008, Yale University School of Nursing. All rights reserved.
Problems or questions: teencope@yale.edu

Figure 3. Managing Diabetes screenshot displaying webpage for nutrition segment.

Healthy Eating

Learning About Nutrition - Fats

Fat is another component in food. Some foods, like fruits and vegetables have almost no fat. Other foods like oil, butter, and beef have plenty of fat.

Dietary fat helps fuel the body, helps absorb some vitamins, and insulates the nervous system.

Dietary fat is an important part of a healthy diet, but some fats are better for you than others. It is best to limit saturated and trans fats.

Good fats:

- Unsaturated fats

Bad fats:

- Saturated fats
- Trans fats

Source: www.kidshealth.org

Sponsored by NIH: R01NR4009.
Copyright © 2008 Yale School of Nursing. All rights reserved.
Problems or questions: teencope@yale.edu

Figure 4. Managing Diabetes screenshot displaying problem-solving questions for glucose control segment.

Procedures

Institutional Review Boards at all clinical sites approved the study. Youth and parents were approached in the clinic setting and informed consent/assent was obtained by trained research personnel. Demographic data were completed by parents at enrollment and email communication was subsequently established with the youth. The youth were sent a link to a password-protected data collection website, and parents were notified of this communication. Internet sites were password protected with all data encrypted and stored on a secure server with hardware and software firewalls. If the youth did not complete online data collection within 1 week, the youth and parents were called approximately 3 times over a 1- to 2-month period as a reminder. Several (2-4) emails and a postcard were also sent during that time in an attempt to re-engage the family. If online data collection was not completed within 3 months, the youth were considered to be passively refusing study participation.

Upon completion of baseline data collection, an automated email was sent to the youths and their parents/guardians to identify their group assignment and provide a link to the appropriate program. (Group assignment was previously determined by randomization of study identification numbers in blocks of 10 to either TEENCOPE or Managing Diabetes.) A unique password was provided to each participant and they were instructed to change this password the first time they logged on to the program. Each program had 5 sessions that were released weekly and took approximately 30 minutes to complete.

A protocol was implemented to enhance participation. Participants were contacted by phone after the first session was released to ensure that they had received the email and were able to access the program. If a youth did not complete a session within 1 week, weekly email, phone, or postcard reminders were sent. Parents were contacted by email if youths did not complete a lesson after 3 weeks. Participants received a gift card (US $25-$30) for completion of online questionnaires evaluating program efficacy.

Data Collection

Data for this analysis included demographic, recruitment, participation, satisfaction data, and depressive symptoms. Demographic data included gender, age, race/ethnicity, and household income. Recruitment was categorized into 3 groups: (1) refusal at point of contact (refused), (2) those who consented but never established email communication with research personnel or did not complete baseline data (passive refusal), and (3) those who participated (enrolled). Data were not available on income and race/ethnicity for youths who refused at the point of contact. Participation in the Internet programs was categorized into 2 groups: participants who completed the goal of at least 80% of sessions (4 of 5) and participants who completed less than 80% of sessions.

Satisfaction was evaluated with a 6-item survey on how helpful, enjoyable, engaging, easy to use, and worthwhile the program was, as well as how much the skills were practiced. Items were rated on a 5-point Likert scale with higher scores indicative of higher satisfaction (1=not at all; 5=very satisfied). Cronbach alpha in this sample was .73.

Metabolic control was measured with glycosylated hemoglobin (A1C), an estimate of the adolescent’s glycemic control over the past 8 to 12 weeks. The American Diabetes Association (ADA) recommendation for children aged 6 to 12 years is less than 8% [5].

Depressive symptoms were measured with the Children’s Depression Inventory (CDI), a self-report inventory for youth with items on mood, vegetative functions, and interpersonal behaviors [30]. It contains 27 multiple-choice items that yield total scores from 0 to 54. The item that addresses suicidal ideation was eliminated because of the inability to respond immediately to a positive endorsement through the Internet. Higher scores reflect a higher number of symptoms. Youth who scored above the threshold for depression (≥12) were contacted by a qualified study staff member (psychologist, nurse, or social worker) who conducted a depression assessment and made appropriate referrals. Cronbach alpha in our sample was .90.

Data Analytic Plan

To test for demographic differences in recruitment and participation, Chi-square analyses were conducted to compare categories of recruitment and participation by gender, age, income level, and race/ethnicity. Participation was also compared by metabolic control (A1C within the recommended range, ≤8% above the recommended range) and depressive symptoms (high CDI score ≥12, normal CDI score <12). To test for demographic differences in satisfaction, *t* tests or 1-way analyses of variance (ANOVA) were conducted.

Results

Description of Sample

A total of 518 youth were initially approached to participate in the study. Eight of those teens participated in the study, but were excluded from analysis due to ineligibility. A total of 320 (62.7%) eligible teens enrolled in the study, 112 (22.0%) refused, and 78 (15.3%) passively refused (Figure 5). Reasons for refusal included no interest (n=63, 56.3%), time (n=19, 17.0%), lack of easy access to Internet (n=8, 7.1%), and unknown (n=22, 19.6%).

The mean age was 12.3 years (SD 1.1) and 177 (55.3%) of the sample were female (see Table 1). A total of 204 (65.2%) were non-Hispanic white and 109 (34.8%) were black, Hispanic, or other. Overall, approximately half of the sample (n=165, 52.5%) were above the recommended range for metabolic control (A1C>8%). In terms of depressive symptoms, 53 (16.6%) scored above the clinical cutoff for depression on the CDI (≥12).

Table 1. Sample demographics of teen participants (N=320).

Characteristic		Participants
Age (years), mean (SD)		12.3 (1.1)
Duration of diabetes (years), mean (SD)		6.1 (3.5)
Gender, n (%)		
	Male	143 (44.7)
	Female	177 (55.3)
Race/ethnicity, n (%) [a]		
	White/non-Hispanic	204 (65.2)
	Black/Hispanic/other	109 (34.8)
Annual household income (US $), n (%) [a]		
	<$40,000	65 (21.0)
	$40,000 - $79,999	87 (28.2)
	>$80,000	157 (50.8)
Insulin therapy, n (%) [a]		
	Pump	189 (59.1)
	Injection (basal)	77 (24.1)
	Injection (conventional)	51 (15.9)
A1C>8%, n (%)[a]		165 (52.5)
CDI≥12, n (%)		53 (16.6)

[a] The totals for these variables do not equal 320 because some participants chose not to answer these questions or data were not available. Percentages represent valid percent.

Recruitment

Chi-square analyses were used to test for demographic differences by recruitment category (see Table 2). There were no significant differences in recruitment category by gender (χ^2_2=4.0, P=.14) or age group (11-12 vs 13-14; χ^2_2=2.3, P=.31). There was, however, a significant difference for race/ethnicity (χ^2_6=34.3, P<.001) with respect to enrollment. White youth were more likely to enroll and less likely to passively refuse, whereas black, Hispanic, or mixed race/ethnicity youth were less likely to enroll and more likely to passively refuse. There was also a significant difference by income (χ^2_4=30.5, P<.001), with teens from the lowest income category (annual household income <US $40,000) less likely to enroll and more likely to passively refuse than teens from the higher-income categories (annual household income >US $40,000). It is important to mention that data were not available on race/ethnicity and income for all recruitment categories; data were unavailable on the race/ethnicity and income of youth who refused at the point of contact.

Table 2. Demographic differences by for participants who enrolled (n=320), those who refused at point of contact (n=112), and those who refused after consent (n=78) of the total 510 eligible youth who were approached to participate.

Characteristic	Enrolled n (%)	Refused at point of contact[a] n (%)	Refused after consent n (%)	Chi-square χ^2 (df)	*P*
Gender					
Male	143 (44.7)	57 (23.8)	40 (16.7)	4.0 (2)	.14
Female	177 (55.3)	46 (17.6)	38 (14.6)		
Total[b]	320 (63.9)	103 (20.6)	78 (15.6)		
Race					
Black	25 (8.0)	—	13 (16.8)	34.3 (6)	<.001
Hispanic	59 (18.8)	—	30 (39.0)		
White	204 (65.2)	—	27 (35.1)		
Biracial or multiracial	25 (8.0)	—	7 (9.1)		
Total[b]	313 (80.3)	—	77 (19.7)		
Household income (US $)					
<$40,000	65 (21.0)	—	35 (35.0)	30.5 (4)	<.001
$40,000-$80,000	87 (28.2)	—	20 (18.7)		
>$80,000	157 (50.8)	—	18 (10.3)		
Total[b]	309 (80.9)	—	73 (19.1)		
Age group					
11-12 years	185 (57.8)	53 (49.5)	45 (15.9)	2.3 (2)	.31
13-14 years	135 (42.2)	54 (50.5)	33 (14.9)		
Total[b]	320 (63.4)	107 (21.2)	78 (15.4)		

[a] Data about race/ethnicity and income not available for this group.

[b] Percentage noted is valid percent, taking into account missing data.

Participation

Participation in the Internet programs was high, with 250 (78.1%) of youth completing at least 4 of 5 sessions, 39 (12.2%) completing 1 to 3 sessions, and 31 (9.7%) completing no sessions. The mean number of sessions completed was 4.08 (SD 1.64) across both groups. There was no significant difference in participation between groups; 129 teens in TEENCOPE participated (completed at least 80% of sessions) at a rate of 77.2%, whereas 121 teens in Managing Diabetes participated at a rate of 79.1% (χ^2_1=0.2, *P*=.69).

Results of the Chi-square analyses to test for demographic differences by participation are provided in Table 3. There was no significant difference by gender (χ^2_1=1.0, *P*=.31), race/ethnicity (χ^2_3=3.1, *P*=.37), or age group (χ^2_1=3.1, *P*=.08). There was a significant difference for income (χ^2_2=12.6, *P*=.002), with those in the lowest income category (annual household income <US $40,000) less likely to participate, and those in the highest income category (annual household income >US $80,000) most likely to participate. There was no significant difference in participation for metabolic control; adolescents who had an A1C below the recommended cutoff (<8%) were no more likely to participate than those above the cutoff (χ^2_1=0.2, *P*=.63). Lastly, depressive symptoms (CDI score) were significantly related to participation (χ^2_1=3.9, *P*=.05); youth who scored above the clinical cutoff on depressive symptoms (≥12) were less likely to complete 4 or more sessions than those who scored in the normal range.

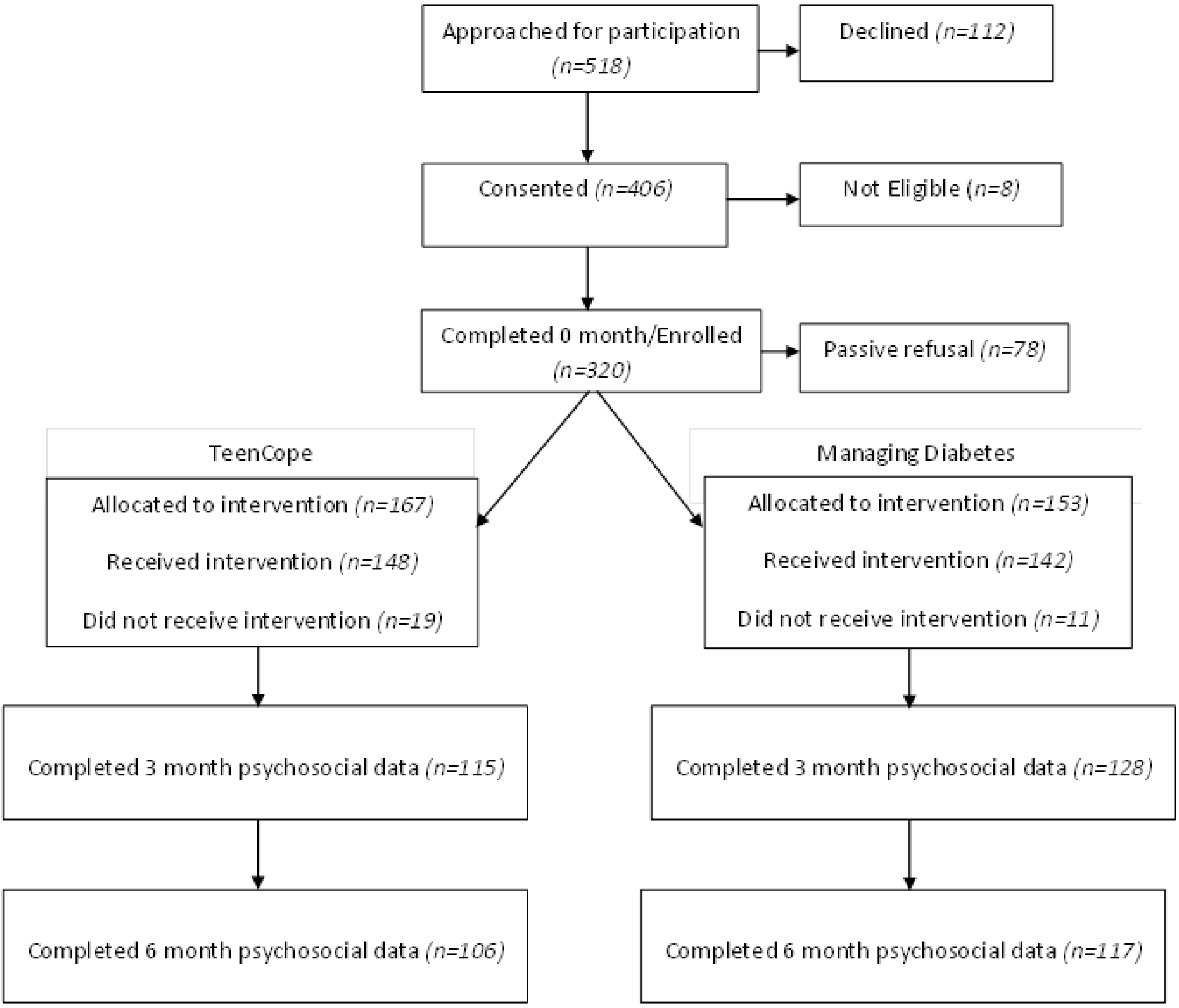
Table 3. Demographic differences in participation for participators (completed at least 4 sessions or 80%) and nonparticipators (completed <80% of sessions).

Characteristic		Participator n (%)[a]	Nonparticipator n (%)[a]	Chi-square χ^2 (df)	*P*
Gender					
	Male	108 (43.2)	35 (50.0)	1.0 (1)	.31
	Female	142 (56.8)	35 (50.0)		
	Total[b]	250 (78.1)	70 (21.9)		
Race					
	Black	20 (8.2)	5 (7.2)	3.1 (3)	.37
	Hispanic	46 (18.9)	13 (18.8)		
	White	162 (66.4)	42 (60.9)		
	Biracial or multiracial	16 (6.6)	9 (13.0)		
	Total[b]	244 (78.0)	69 (22.0)		
Age group					
	11-12 years	151 (60.4)	34 (48.6)	3.1 (1)	.08
	13-14 years	99 (39.6)	36 (51.4)		
	Total[b]	250 (78.1)	70 (21.9)		
Household income (US $)					
	<$40,000	41 (17.1)	24 (34.8)	12.6 (2)	.002
	$40,000-$80,000	66 (27.4)	21 (30.4)		
	>$80,000	133 (55.4)	24 (34.8)		
	Total[b]	240 (77.7)	69 (22.3)		
A1C					
	Recommended (<8%)	118 (48.2	31 (44.9)	0.2 (1)	.63
	High (>12%)	127 (51.8)	38 (55.1)		
	Total[b]	245(78.0)	69 (22.0)		
CDI					
	Normal (<12)	214 (85.6)	53 (75.7)	3.9 (1)	.05
	High (>12)	36 (14.4)	17 (24.3)		
	Total[b]	250 (78.1)	70 (21.9)		

[a] The total number is different by category, due to missing data; some participants chose not to answer questions about race/ethnicity and income.

[b] Total percentages are valid percent, accounting for missing data.

Figure 5. Consort flow diagram.



Satisfaction

Satisfaction was high with both programs, with no significant difference between groups. The mean satisfaction score was 3.97 (SD 0.71) for TEENCOPE and 3.89 (SD 0.56) for Managing Diabetes. There were significant gender, race/ethnicity, and income differences, in that girls (t_{202}=3.28, P=.001), black, Hispanic, or mixed race/ethnicity (t_{196}=2.42, P=.02), and low-income youth ($F_{2,201}$=3.80, P=.02) reported higher satisfaction. There was no difference in satisfaction by age or depressive symptoms.

Summary

Results of the analysis are summarized in Table 4. Lower enrollment rates were demonstrated in youth with Type 1 diabetes who were black, Hispanic, or of mixed race/ethnicity, and of lower income. Lower participation was demonstrated in youth from low-income families. Higher satisfaction was reported by girls, lower-income youth, and black, Hispanic, or mixed race/ethnicity youth.

Table 4. Statistically significant comparisons of recruitment category, participation, and satisfaction by demographic characteristics.

Outcome	*P* values			
	Gender	Age	Race/ethnicity	Income
Recruitment/enrollment	—	—	<.001	<.001
Participation	—	—	—	.002
Satisfaction	<.001	—	.02	.02

Discussion

The 2 psychoeducational eHealth programs for youth with Type 1 diabetes in our study were able to reach a diverse sample, reflective of national prevalence estimates. Overall, there was high participation and satisfaction with the 2 programs. Given the huge influx of eHealth interventions designed for pediatric populations, it is important to understand who they are likely to reach and benefit. Results from the current study highlight demographic differences in recruitment/enrollment, participation, and satisfaction with psychosocial and educational eHealth programs for youth with Type 1 diabetes. Results indicate that black, Hispanic, or mixed race/ethnicity youth with Type 1 diabetes are less likely to enroll in Internet-based research than white youth; thus, creative recruitment approaches are needed. Lower-income youth were also less likely to participate than higher-income youth, possibly due to problems with access. However, once enrolled, black, Hispanic, or mixed race/ethnicity youth and lower-income youth with Type 1 diabetes were as highly satisfied or more satisfied with the eHealth programs as white youth and higher-income youth. These results support the idea that eHealth programs designed for pediatric populations have the potential to reach diverse youth and be appealing to them.

Overall, the rate of enrollment in the study (63%) was average for youth psychoeducational research, which ranges from 49% to 73% [10,21]. The majority of youth who declined participation at the point of contact indicated that they were not interested in the study. As we have shown in a previous study, youth and their families may be less likely to participate in a preventive intervention of this type before the onset of puberty when poorer metabolic control is common [31]. Very few youth indicated that lack of Internet access was a reason for not participating, but they may not have wanted to disclose their lack of Internet access to the clinical or research personnel.

There were no differences in recruitment category with respect to gender or age. There were, however, significant demographic differences in enrollment, with black, Hispanic, or mixed race/ethnicity and lower-income youth more likely to passively refuse, and white youth more likely to enroll. It is possible that black, Hispanic, or mixed race/ethnicity and lower-income youth consent to take part in studies because they feel pressure to please a respected individual (eg, health care providers, research staff) [32], rather than out of a desire to participate. Alternatively, given that black, Hispanic, or mixed race/ethnicity and low-income youth are less likely to be online than white youth [13], youth who passively refused may have had problems with Internet access. In addition, some of the families who had Internet service at the time of enrollment may have later lost or canceled the service because of financial difficulties. These problems with passive refusal (ie, youth who consented to participate but never completed baseline data) are similar to another Internet study with a pediatric population that also was unable to contact some youth after initial consent [33]. Rates of passive refusal in other studies evaluating an eHealth program in youth have ranged from 13% to 40% [18,34-36].

The pattern of results for participation is somewhat different from enrollment; there were no significant differences in race/ethnicity for participation, but there was a significant difference by income. Youth in the lowest income category (family income <US $40,000 year) were significantly less likely to participate than those in the higher-income categories. Again, lower levels of participation may reflect problems with Internet access. Previous studies have shown that lower-income families are less likely to have a home computer and home Internet access [20]. Similarly, lower-income youth report going online less often than higher-income youth; 39% of lower-income teens go online daily vs 75% of higher-income teens [37]. Further, teens from lower-income families are more likely to use the Internet at school, whereas 99% of those from higher-income families access the Internet at home [37]. Thus, it may be less convenient for lower-income teens to participate in eHealth programs than higher-income teens.

Providing options for participation in eHealth programs at schools and clinics may enhance participation. Yet, none of the participants in this study took advantage of Internet access at the clinic for program participation, despite it being offered. Development of eHealth programs for youth in the future may need to use multiple platforms, such as the Internet and smartphones. Currently, 75% of adolescents have cell phones [37]. Latino and black youth are more likely than white youth to access the Internet by cell phone [37]; thus, eHealth programs that are able to be viewed on both the Internet and smartphones may reach more youth of diverse races/ethnicities.

A number of eHealth programs provided on mobile phones with text messaging or smartphone applications have been developed for youth with Type 1 diabetes to enhance blood glucose monitoring [34,38,39] and/or diabetes treatment [40-42]. Improvements in blood glucose monitoring [34] and adherence [41] have been demonstrated. Metabolic control improved only in programs that provided additional components, such as intensive diabetes treatment [41] or behavioral contracting [38]. Studies have primarily been small, 1-group pilot studies, and participation and satisfaction have not been consistently reported. When reported, the reach of the program to youth of diverse races/ethnicities has been low [41], participation has been variable [34,40], and satisfaction has been high [34,40].

Limited literature is available on strategies to promote greater engagement of youth in behavioral interventions, with even less information on the use of the Internet and mobile technologies for minority and low-income youth. Designers of eHealth programs have identified that involving targeted users in the design and development of programs is critical to enhancing user experience and acceptability [43]. It has been proposed that 80% of the impact of an eHealth program is determined during the design phase [44]. Thus, including youth of diverse races and ethnicities in the design and development of eHealth programs appears critical. The use of social media has also been successful in engaging diverse youth in health promotion content by using interactive blogging, connection with others, and creative expression [45]. From the perspectives of African American youth, some evidence supports the value of parental modeling and social support to encourage participation in certain behavioral interventions [46]. Parents of black, Hispanic, or mixed race/ethnicity youth may be key to enlisting their young adolescents' interest in future eHealth programs. Researchers need to understand parental views on the advantages, safety, and potential health benefits to having their adolescents participate. Novel approaches for obtaining information on parental level of enthusiasm regarding a proposed eHealth investigation through in-person or Skype focus groups could aid in the design of future investigations with minority groups.

Interestingly, our results on satisfaction indicated that black, Hispanic, or mixed race/ethnicity youth and lower-income youth were more satisfied with the program than higher-income and white youth. These high levels of satisfaction are likely a reflection of the inclusion of a culturally diverse "cast" of characters in the program and developmentally appropriate topics and examples. Addressing cultural issues has been shown to foster adherence [47]. Further, we included input from teens at every stage of the program development and paid careful attention to the usability of the program [48]. Finally, we kept text to a minimum, which is more appealing to youth, especially those with low health literacy [49]. Girls were also more satisfied than boys, which may reflect their interest in behavioral topics (ie, interpersonal skills). It has also been hypothesized that girls may respond more to a cognitive-behavioral program based on social learning theory [50].

Finally, there were differences in participation related to depressive symptoms, such that teens who scored above the clinical cutoff were less likely to participate than those who scored below the cutoff. The presence of depressive symptoms may negatively impact motivation for many youth, which has been shown in a previous study of an eHealth smoking cessation program [27]. Youth with higher depressive symptoms may be more likely to participate in programs that they believe will help with their depression [28].

Limitations

This study provides important new information regarding the role of demographic factors in the study of eHealth programs, but there are several limitations. First, it is important to acknowledge that many youth in our study needed prompts and reminders from research staff to achieve a high level of participation. Further, we did not have data on race/ethnicity and household income for the families who did not enroll in the study. Finally, although the sample was diverse, race/ethnicity and income were highly correlated; only 23% of the high-income youth were not white, making it difficult to determine the relative effect of race/ethnicity and income on participation and satisfaction.

Directions for Future Research

With the move toward more use of eHealth interventions, it is increasingly important to design programs to maximize recruitment and retention. Although findings from the current study provide important information regarding the demographic factors that are important to consider, future studies are still needed to tease apart the different effects of race/ethnicity and income. Better reporting on recruitment yield (including number approached, ineligible, refused at point of contact, and passively refused) and demographic characteristics of participants and nonparticipants of eHealth programs is needed [51]. In addition, more consistent reporting of youth participation and satisfaction in eHealth research, as well as reasons for refusal are indicated. Recently, a CONSORT eHealth checklist has been published which identifies intended dose and actual dose as well as participant usage over time as important data to report in eHealth clinical trials [52]. Community participatory research strategies to elicit personal perspectives of parents and teens from diverse economic and racial/ethnic backgrounds are needed when designing future eHealth technologies to promote diabetes management. Finally, research on factors that influence youth participation and satisfaction with eHealth programs, including platform for Internet access (ie, home vs smartphone), is also needed.

Conclusion

As an innovative approach, the use of eHealth programs can improve access to psychoeducational programs for youth of diverse races/ethnicities, socioeconomic status, and with varying chronic illnesses. It is critical that eHealth interventions reach the targeted population to maximize external validity and generalizability of study results. More evaluation of the recruitment, participation, retention, and satisfaction of youth of diverse race/ethnicity and their parents is essential and needed for wider dissemination of future eHealth research.

Acknowledgments

This research was supported by grant number R01-NR04009 from the National Institute of Nursing Research (principal investigators Grey and Whittemore). Special acknowledgment to Heather Jacobs, RN, MPH; Caitlin Paul, BA; and Lauren Liberti, MS, for their assistance in preparation of this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Consort e-health checklist [52].

[PDF File (Adobe PDF File), 52KB - jmir_v15i1e15_app1.pdf]

References

1. SEARCH for Diabetes in Youth Study Group, Liese AD, D'Agostino RB, Hamman RF, Kilgo PD, Lawrence JM, et al. The burden of diabetes mellitus among US youth: prevalence estimates from the SEARCH for Diabetes in Youth Study. Pediatrics 2006 Oct;118(4):1510-1518 [FREE Full text] [doi: 10.1542/peds.2006-0690] [Medline: 17015542]
2. Bell RA, Mayer-Davis EJ, Beyer JW, D'Agostino RB, Lawrence JM, Linder B, SEARCH for Diabetes in Youth Study Group. Diabetes in non-Hispanic white youth: prevalence, incidence, and clinical characteristics: the SEARCH for Diabetes in Youth Study. Diabetes Care 2009 Mar;32 Suppl 2:S102-S111 [FREE Full text] [doi: 10.2337/dc09-S202] [Medline: 19246575]
3. Lawrence JM, Mayer-Davis EJ, Reynolds K, Beyer J, Pettitt DJ, D'Agostino RB, SEARCH for Diabetes in Youth Study Group. Diabetes in Hispanic American youth: prevalence, incidence, demographics, and clinical characteristics: the SEARCH for Diabetes in Youth Study. Diabetes Care 2009 Mar;32 Suppl 2:S123-S132 [FREE Full text] [doi: 10.2337/dc09-S204] [Medline: 19246577]
4. Mayer-Davis EJ, Beyer J, Bell RA, Dabelea D, D'Agostino R, Imperatore G, SEARCH for Diabetes in Youth Study Group. Diabetes in African American youth: prevalence, incidence, and clinical characteristics: the SEARCH for Diabetes in Youth Study. Diabetes Care 2009 Mar;32 Suppl 2:S112-S122 [FREE Full text] [doi: 10.2337/dc09-S203] [Medline: 19246576]
5. American Diabetes Association. Standards of medical care in diabetes--2012. Diabetes Care 2012 Jan;35 Suppl 1:S11-S63. [doi: 10.2337/dc12-s011] [Medline: 22187469]
6. Anderson BJ, Vangsness L, Connell A, Butler D, Goebel-Fabbri A, Laffel LM. Family conflict, adherence, and glycaemic control in youth with short duration Type 1 diabetes. Diabet Med 2002 Aug;19(8):635-642. [Medline: 12147143]
7. Anderson BJ, Holmbeck G, Iannotti RJ, McKay SV, Lochrie A, Volkening LK, et al. Dyadic measures of the parent-child relationship during the transition to adolescence and glycemic control in children with type 1 diabetes. Fam Syst Health 2009 Jun;27(2):141-152 [FREE Full text] [doi: 10.1037/a0015759] [Medline: 19630455]
8. Ingerski LM, Anderson BJ, Dolan LM, Hood KK. Blood glucose monitoring and glycemic control in adolescence: contribution of diabetes-specific responsibility and family conflict. J Adolesc Health 2010 Aug;47(2):191-197 [FREE Full text] [doi: 10.1016/j.jadohealth.2010.01.012] [Medline: 20638012]
9. Anderson BJ, Brackett J, Ho J, Laffel LM. An office-based intervention to maintain parent-adolescent teamwork in diabetes management. Impact on parent involvement, family conflict, and subsequent glycemic control. Diabetes Care 1999 May;22(5):713-721 [FREE Full text] [Medline: 10332671]
10. Grey M, Boland EA, Davidson M, Li J, Tamborlane WV. Coping skills training for youth with diabetes mellitus has long-lasting effects on metabolic control and quality of life. J Pediatr 2000 Jul;137(1):107-113. [doi: 10.1067/mpd.2000.106568] [Medline: 10891831]
11. Wysocki T, Harris MA, Buckloh LM, Mertlich D, Lochrie AS, Mauras N, et al. Randomized trial of behavioral family systems therapy for diabetes: maintenance of effects on diabetes outcomes in adolescents. Diabetes Care 2007 Mar;30(3):555-560. [doi: 10.2337/dc06-1613] [Medline: 17327320]
12. Barlow JH, Ellard DR. Psycho-educational interventions for children with chronic disease, parents and siblings: an overview of the research evidence base. Child Care Health Dev 2004 Nov;30(6):637-645. [doi: 10.1111/j.1365-2214.2004.00474.x] [Medline: 15527474]
13. Pew Internet. Trend data for teens. Washington, DC: Pew Internet & American Life Project; 2009. URL: http://pewinternet.org/Static-Pages/Trend-Data-(Teens).aspx [accessed 2012-12-19] [WebCite Cache ID 6D24tqmBK]
14. Stinson J, Wilson R, Gill N, Yamada J, Holt J. A systematic review of internet-based self-management interventions for youth with health conditions. J Pediatr Psychol 2009 Jun;34(5):495-510 [FREE Full text] [doi: 10.1093/jpepsy/jsn115] [Medline: 19029142]
15. Ritterband LM, Gonder-Frederick LA, Cox DJ. Internet interventions: in review, in use, and into the future. Professional Psychology: Research and Practice 2003;34(5):527-534. [doi: 10.1037/0735-7028.34.5.527]

16. Mulvaney SA, Rothman RL, Wallston KA, Lybarger C, Dietrich MS. An internet-based program to improve self-management in adolescents with type 1 diabetes. Diabetes Care 2010 Mar;33(3):602-604 [FREE Full text] [doi: 10.2337/dc09-1881] [Medline: 20032275]
17. Whittemore R, Jeon S, Jaser SS. Abstracts of the 37th Annual Meeting of the International Society for Pediatric and Adolescent Diabetes (ISPAD). October 19-22, 2011. Miami Beach, Florida, USA. Pediatr Diabetes 2011 Oct;12 Suppl 15(suppl 15):1-152. [doi: 10.1111/j.1399-5448.2011.01816.x] [Medline: 21978093]
18. Mulvaney SA, Rothman RL, Osborn CY, Lybarger C, Dietrich MS, Wallston KA. Self-management problem solving for adolescents with type 1 diabetes: intervention processes associated with an Internet program. Patient Educ Couns 2011 Nov;85(2):140-142 [FREE Full text] [doi: 10.1016/j.pec.2010.09.018] [Medline: 21030194]
19. Sun P, Unger JB, Palmer PH, Gallaher P, Chou CP, Baezconde-Garbanati L, et al. Internet accessibility and usage among urban adolescents in Southern California: implications for web-based health research. Cyberpsychol Behav 2005 Oct;8(5):441-453. [doi: 10.1089/cpb.2005.8.441] [Medline: 16232037]
20. Carroll AE, Rivara FP, Ebel B, Zimmerman FJ, Christakis DA. Household computer and Internet access: The digital divide in a pediatric clinic population. In: AMIA Annu Symp Proc. 2005 Presented at: AMIA Annu Symp Proc; 2005 p. 111-115 URL: http://europepmc.org/abstract/MED/16779012
21. Diviak KR, Wahl SK, O'Keefe JJ, Mermelstein RJ, Flay BR. Recruitment and retention of adolescents in a smoking trajectory study: who participates and lessons learned. Subst Use Misuse 2006;41(2):175-182. [doi: 10.1080/10826080500391704] [Medline: 16393741]
22. Hinshaw SP, Hoagwood K, Jensen PS, Kratochvil C, Bickman L, Clarke G, et al. AACAP 2001 research forum: challenges and recommendations regarding recruitment and retention of participants in research investigations. J Am Acad Child Adolesc Psychiatry 2004 Aug;43(8):1037-1045. [doi: 10.1097/01.chi.0000129222.89433.3d] [Medline: 15266200]
23. Chan DS, Callahan CW, Hatch-Pigott VB, Lawless A, Proffitt HL, Manning NE, et al. Internet-based home monitoring and education of children with asthma is comparable to ideal office-based care: results of a 1-year asthma in-home monitoring trial. Pediatrics 2007 Mar;119(3):569-578 [FREE Full text] [doi: 10.1542/peds.2006-1884] [Medline: 17332210]
24. O'Kearney R, Kang K, Christensen H, Griffiths K. A controlled trial of a school-based Internet program for reducing depressive symptoms in adolescent girls. Depress Anxiety 2009;26(1):65-72. [doi: 10.1002/da.20507] [Medline: 18828141]
25. Williamson DA, Walden HM, White MA, York-Crowe E, Newton RL, Alfonso A, et al. Two-year internet-based randomized controlled trial for weight loss in African-American girls. Obesity (Silver Spring) 2006 Jul;14(7):1231-1243 [FREE Full text] [doi: 10.1038/oby.2006.140] [Medline: 16899804]
26. Frenn M, Malin S, Brown RL, Greer Y, Fox J, Greer J, et al. Changing the tide: an Internet/video exercise and low-fat diet intervention with middle-school students. Appl Nurs Res 2005 Feb;18(1):13-21. [doi: 10.1016/j.apnr.2004.04.003] [Medline: 15812731]
27. Neil AL, Batterham P, Christensen H, Bennett K, Griffiths KM. Predictors of adherence by adolescents to a cognitive behavior therapy website in school and community-based settings. J Med Internet Res 2009;11(1):e6 [FREE Full text] [doi: 10.2196/jmir.1050] [Medline: 19275982]
28. Christensen H, Reynolds J, Griffiths KM. The use of e-health applications for anxiety and depression in young people: challenges and solutions. Early Interv Psychiatry 2011 Feb;5 Suppl 1:58-62. [doi: 10.1111/j.1751-7893.2010.00242.x] [Medline: 21208393]
29. Grey M, Whittemore R, Liberti L, Delamater A, Murphy K, Faulkner MS. A comparison of two internet programs for adolescents with type 1 diabetes: design and methods. Contemp Clin Trials 2012 Jul;33(4):769-776. [doi: 10.1016/j.cct.2012.03.012] [Medline: 22484337]
30. Kovacs M. The Children's Depression, Inventory (CDI). Psychopharmacol Bull 1985;21(4):995-998. [Medline: 4089116]
31. Grey M, Whittemore R, Jaser S, Ambrosino J, Lindemann E, Liberti L, et al. Effects of coping skills training in school-age children with type 1 diabetes. Res Nurs Health 2009 Aug;32(4):405-418 [FREE Full text] [doi: 10.1002/nur.20336] [Medline: 19488997]
32. Thompson D, Canada A, Bhatt R, Davis J, Plesko L, Baranowski T, et al. eHealth recruitment challenges. Eval Program Plann 2006 Nov;29(4):433-440. [doi: 10.1016/j.evalprogplan.2006.08.004] [Medline: 17950873]
33. Ritterband LM, Borowitz S, Cox DJ, Kovatchev B, Walker LS, Lucas V, et al. Using the internet to provide information prescriptions. Pediatrics 2005 Nov;116(5):e643-e647 [FREE Full text] [doi: 10.1542/peds.2005-0404] [Medline: 16263978]
34. Cafazzo JA, Casselman M, Hamming N, Katzman DK, Palmert MR. Design of an mHealth app for the self-management of adolescent type 1 diabetes: a pilot study. J Med Internet Res 2012;14(3):e70 [FREE Full text] [doi: 10.2196/jmir.2058] [Medline: 22564332]
35. Doyle AC, Goldschmidt A, Huang C, Winzelberg AJ, Taylor CB, Wilfley DE. Reduction of overweight and eating disorder symptoms via the Internet in adolescents: a randomized controlled trial. J Adolesc Health 2008 Aug;43(2):172-179 [FREE Full text] [doi: 10.1016/j.jadohealth.2008.01.011] [Medline: 18639791]
36. March S, Spence SH, Donovan CL. The efficacy of an internet-based cognitive-behavioral therapy intervention for child anxiety disorders. J Pediatr Psychol 2009 Jun;34(5):474-487 [FREE Full text] [doi: 10.1093/jpepsy/jsn099] [Medline: 18794187]

37. Lenhart A, Arafeh S, Smith A, Macgill A. The lives of teens and their technology. Washington, DC: Pew Internet & American Life Project; 2012 Apr 24. URL: http://www.pewinternet.org/Reports/2008/Writing-Technology-and-Teens/04-The-Lives-of-Teens-and-Their-Technology/01-Introduction.aspx [accessed 2012-12-19] [WebCite Cache ID 6D24xILOn]
38. Carroll AE, DiMeglio LA, Stein S, Marrero DG. Contracting and monitoring relationships for adolescents with type 1 diabetes: a pilot study. Diabetes Technol Ther 2011 May;13(5):543-549 [FREE Full text] [doi: 10.1089/dia.2010.0181] [Medline: 21406011]
39. Carroll AE, DiMeglio LA, Stein S, Marrero DG. Using a cell phone-based glucose monitoring system for adolescent diabetes management. Diabetes Educ 2011;37(1):59-66. [doi: 10.1177/0145721710387163] [Medline: 21106908]
40. Franklin VL, Greene A, Waller A, Greene SA, Pagliari C. Patients' engagement with "Sweet Talk" - a text messaging support system for young people with diabetes. J Med Internet Res 2008;10(2):e20 [FREE Full text] [doi: 10.2196/jmir.962] [Medline: 18653444]
41. Franklin VL, Waller A, Pagliari C, Greene SA. A randomized controlled trial of Sweet Talk, a text-messaging system to support young people with diabetes. Diabet Med 2006 Dec;23(12):1332-1338. [doi: 10.1111/j.1464-5491.2006.01989.x] [Medline: 17116184]
42. Frøisland DH, Arsand E, Skårderud F. Improving diabetes care for young people with type 1 diabetes through visual learning on mobile phones: mixed-methods study. J Med Internet Res 2012;14(4):e111 [FREE Full text] [doi: 10.2196/jmir.2155] [Medline: 22868871]
43. Skinner HA, Maley O, Norman CD. Developing internet-based eHealth promotion programs: the Spiral Technology Action Research (STAR) model. Health Promot Pract 2006 Oct;7(4):406-417. [doi: 10.1177/1524839905278889] [Medline: 16840770]
44. Thackara J. In the Bubble: Designing in a Complex World. Cambridge, MA: MIT Press; 2005.
45. Norman CD, Skinner HA. Engaging youth in e-health promotion: lessons learned from a decade of TeenNet research. Adolesc Med State Art Rev 2007 Aug;18(2):357-69, xii. [Medline: 18605651]
46. Wright MS, Wilson DK, Griffin S, Evans A. A qualitative study of parental modeling and social support for physical activity in underserved adolescents. Health Educ Res 2010 Apr;25(2):224-232 [FREE Full text] [doi: 10.1093/her/cyn043] [Medline: 18703530]
47. Cashen MS, Dykes P, Gerber B. eHealth technology and Internet resources: barriers for vulnerable populations. J Cardiovasc Nurs 2004;19(3):209-14; quiz 215. [Medline: 15191264]
48. Whittemore R, Grey M, Lindemann E, Ambrosino J, Jaser S. Development of an Internet coping skills training program for teenagers with type 1 diabetes. Comput Inform Nurs 2010;28(2):103-111 [FREE Full text] [doi: 10.1097/NCN.0b013e3181cd8199] [Medline: 20182161]
49. Kutner M, Greenberg E, Jin Y, Paulson C. The Health Literacy of America's Adults: Results From the 2003 National Assessment of Adult Literacy. Washington, DC: National Center for Education Statistics; 2006 Sep. URL: http://nces.ed.gov/pubs2006/2006483.pdf [accessed 2013-01-17] [WebCite Cache ID 6Dk94tWZo]
50. Mauriello LM, Ciavatta MM, Paiva AL, Sherman KJ, Castle PH, Johnson JL, et al. Results of a multi-media multiple behavior obesity prevention program for adolescents. Prev Med 2010 Dec;51(6):451-456 [FREE Full text] [doi: 10.1016/j.ypmed.2010.08.004] [Medline: 20800079]
51. Glasgow RE, Strycker LA, Kurz D, Faber A, Bell H, Dickman JM, et al. Recruitment for an internet-based diabetes self-management program: scientific and ethical implications. Ann Behav Med 2010 Aug;40(1):40-48. [doi: 10.1007/s12160-010-9189-1] [Medline: 20411443]
52. 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

A1C: glycosylated hemoglobin
ADA: American Diabetes Association
CDI: Children’s Depression Inventory

Edited by G Eysenbach; submitted 10.05.12; peer-reviewed by M Frenn; comments to author 25.08.12; revised version received 10.09.12; accepted 02.10.12; published 29.01.13.

Please cite as:
Whittemore R, Jaser SS, Faulkner MS, Murphy K, Delamater A, Grey M, TEENCOPE Research Group
Type 1 Diabetes eHealth Psychoeducation: Youth Recruitment, Participation, and Satisfaction
J Med Internet Res 2013;15(1):e15
URL: http://www.jmir.org/2013/1/e15/
doi:10.2196/jmir.2170
PMID:23360729

Original Paper

Periodic Email Prompts to Re-Use an Internet-Delivered Computer-Tailored Lifestyle Program: Influence of Prompt Content and Timing

Francine Schneider[1], MSc; Hein de Vries[1], PhD; Math Candel[2], PhD; Angelique van de Kar[3], PhD; Liesbeth van Osch[1], PhD

[1]CAPHRI, Department of Health Promotion, Maastricht University, Maastricht, Netherlands
[2]CAPHRI, Department of Methodology and Statistics, Maastricht University, Maastricht, Netherlands
[3]Veiligheidsregio Limburg-Noord, Nijmegen, Netherlands

Corresponding Author:
Francine Schneider, MSc
CAPHRI
Department of Health Promotion
Maastricht University
P.O. Box 616
Maastricht, 6200 MD
Netherlands
Phone: 31 433882827
Fax: 31 433671032
Email: francine.schneider@maastrichtuniversity.nl

Abstract

Background: Adherence to Internet-delivered lifestyle interventions using multiple tailoring is suboptimal. Therefore, it is essential to invest in proactive strategies, such as periodic email prompts, to boost re-use of the intervention.

Objective: This study investigated the influence of content and timing of a single email prompt on re-use of an Internet-delivered computer-tailored (CT) lifestyle program.

Methods: A sample of municipality employees was invited to participate in the program. All participants who decided to use the program received an email prompting them to revisit the program. A 2×3 (content × timing) design was used to test manipulations of prompt content and timing. Depending on the study group participants were randomly assigned to, they received either a prompt containing standard content (an invitation to revisit the program), or standard content plus a preview of new content placed on the program website. Participants received this prompt after 2, 4, or 6 weeks. In addition to these 6 experimental conditions, a control condition was included consisting of participants who did not receive an additional email prompt. Clicks on the uniform resource locator (URL) provided in the prompt and log-ins to the CT program were objectively monitored. Logistic regression analyses were conducted to determine whether prompt content and/or prompt timing predicted clicking on the URL and logging in to the CT program.

Results: Of all program users (N=240), 206 participants received a subsequent email prompting them to revisit the program. A total of 53 participants (25.7%) who received a prompt reacted to this prompt by clicking on the URL, and 25 participants (12.1%) actually logged in to the program. There was a main effect of prompt timing; participants receiving an email prompt 2 weeks after their first visit clicked on the URL significantly more often compared with participants that received the prompt after 4 weeks (odds ratio [OR] 3.069, 95% CI 1.392-6.765, *P*=.005) and after 6 weeks (OR 4.471, 95% CI 1.909-10.471, *P*=.001). Furthermore, participants who received an email prompt 2 weeks after their first visit logged in to the program significantly more often compared to participants receiving the prompt after 6 weeks (OR 16.356, 95% CI 2.071-129.196, *P*=.008). A trend was observed with regard to prompt content. Participants receiving a prompt with additional content were more likely to log in to the program compared to participants who received a standard prompt. However, this result was not statistically significant (OR 2.286, 95% CI 0.892-5.856, *P*=.09).

Conclusions: The key findings suggest that boosting revisits to a CT program benefits most from relatively short prompt timing. Furthermore, a preview of new website content may be added to a standard prompt to further increase its effectiveness in persuading people to log in to the program.

(J Med Internet Res 2013;15(1):e23) doi:10.2196/jmir.2151

KEYWORDS

Computer tailoring; Internet-delivered intervention; intervention adherence; periodic email prompts; prompt content; prompt timing

Introduction

Internet-delivered lifestyle interventions applying computer-tailoring techniques [1-3] have reported positive effects for multiple health behaviors, such as physical activity [4,5], fruit and vegetable intake [6,7], smoking cessation [8-11], and alcohol consumption [12,13]. Furthermore, providing computer-tailored (CT) advice on multiple occasions (multiple tailoring) has proven to significantly add to their impact [7,14]. However, despite promising prospects, actual exposure to these interventions remains limited [15]. Exposure not only refers to the level of first-time use of the intervention, but also to the quality and quantity of intervention use [16,17]. A small proportion of the potential target population actually accesses the intervention [16,18,19], and the level of adherence to these interventions is even lower [20], making attrition a common and urgent problem in Internet-based trials [15,21]. Non-usage attrition, in which participants lose interest in the intervention and refrain from continued use, is hindering actual impact on public health [17,21-25].

Research on non-usage attrition has demonstrated an initial rapid decline in program use over the first few weeks [15,21]. During these weeks, participants lose interest in the program or realize that it does not meet their wishes or expectations. However, achieving health behavior change is a complex and lengthy process requiring continuous guidance to maximize intervention effects [22,26]. Therefore, repeated use of interventions using multiple tailoring must be stimulated to allow processing of the entire intervention content and engagement in its effective components [21,27-29]. Furthermore, re-use is important because it offers visitors the opportunity to self-monitor their level of behavior change and receive additional personalized advice regarding strategies to increase or maintain their current level of behavior change [30]. It is essential to minimize non-usage attrition at an early stage by investing in strategies that boost revisits to Internet-delivered interventions using multiple tailoring.

A substantial amount of Internet-delivered interventions use reactive strategies to achieve re-use, implying that a passive approach is used in which users themselves must undertake action to repeatedly benefit from the intervention content [31]. However, because preventing non-usage attrition is a very strenuous process [21,25,32], efforts should be put into attaining loyalty to the intervention directly after initiation by utilizing more proactive strategies [17]. The use of periodic email prompts has been proposed as an effective proactive strategy to boost re-use of interventions [33]. Although the number of interventions employing proactive strategies is increasing, most studies merely explored the efficacy of the whole intervention, including prompting, instead of focusing on the added value of periodic prompting as a strategy to boost re-use. Furthermore, those studies indicating that sending periodic email prompts significantly increased program re-use [17,33] recommended further examination and refinement to maximize their potential [17].

More specific evidence on the positive effects of the use of email to promote direct action stems from the field of e-marketing [34,35]. Within this field, the use of email advertisements is rapidly growing because of the increase in email users and their potential to reach large numbers of people at relatively low cost and effort. In addition, in the field of health promotion, emails have been used to prevent attrition in Web-based trials. Several studies have indicated that sending email reminders is an effective strategy to increase response rates in online data collection [36-38]. Furthermore, within one of our previous studies that examined the effect of using periodic email prompts [17], a positive, however modest, effect on re-use of the program was found. Within that study, an email prompt was sent 3 months after a first visit to the intervention. This modest effect might imply that a 3-month period is too long, causing people to forget about the program and their participation, which is in line with previous studies pointing out an elevated non-usage attrition level at the beginning of the intervention period [15,21]. Therefore, we recommended future research to focus on testing strategies to optimize the effect of email prompts [17]. It is imperative to investigate the effect of using shorter prompting intervals and to determine the optimal interval at which prompts should be sent. In contrast, this modest effect of a periodic email prompt might also imply that prompt content is suboptimal. Since people tend to disengage from the intervention relatively shortly after intervention initiation because of a reduced levels of interest [21], involvement in the intervention content and subsequent sessions is likely to decrease. According to the elaboration likelihood model (ELM), people with a low level of involvement are less likely to process arguments used [39]. As a result, argument-based persuasion techniques used in an email prompt persuading participants to re-use the intervention might be insufficiently processed. To increase persuasiveness of the email prompts for people with a low level of involvement in the issue, the current study tested the effect of adding a peripheral cue to a standard email prompt [39]. This peripheral cue consisted of the addition of a preview of new website content. This preview served as a teaser to increase curiosity for the remaining website content [40]. Instead of basing a decision to re-use the program on the argument posed in the email, participants might simply react out of sheer curiosity [39]. It was hypothesized that a prompt with additional content was more effective in persuading people to revisit the CT lifestyle program compared to a standard prompt.

This study aimed to assess the added value of periodic email prompts to boost revisits to an Internet-delivered CT lifestyle program. We aimed to answer 2 questions: (1) which prompt timing interval is most effective in boosting re-use of the Internet-delivered CT lifestyle program and (2) which prompt content is most effective in boosting re-use of the program? We addressed these questions among participants in an Internet-delivered CT intervention aimed at multiple health behaviors: increasing physical activity, increasing fruit and vegetable intake, smoking cessation, and decreasing alcohol consumption.

Methods

Procedure and Participants

This study was conducted in close collaboration with the Regional Public Health Service (RPHS) in the northern part of the Dutch region, Limburg. All people employed by the municipalities in this region were invited by the RPHS to participate in an Internet-delivered CT lifestyle program developed by our research group [41,42]. Invitations for this program were placed on the intranet of all municipalities and were also directly emailed to employees. The invitation contained a uniform resource locator (URL) that directed interested participants directly to the program. The program provided employees the opportunity to receive free-of-charge CT feedback about their current health behavior (physical activity, fruit and vegetable intake, alcohol consumption, and smoking) and assistance in changing these health behaviors.

Participants who logged in to the program to obtain CT advice were sent an email prompting them to re-use the program. Re-use was encouraged to allow participants to monitor their behavior change and to obtain iterative CT advice aimed at behavior change and relapse prevention. Furthermore, participants were offered an opportunity to obtain CT advice regarding an additional health behavior. Finally, re-use was stimulated to keep participants informed about new content added to the program website.

Design

A 2×3 design was used to test the effect of 2 factors: prompt content (standard and standard+) and prompt timing (2, 4, and 6 weeks). The standard prompt (SP) contained a message reminding people about their previous visit to the program and invited them to re-use the program to monitor their progress and obtain additional feedback. The second version of the prompt (SP+) contained standard content complemented with a message alerting people to new content added to the intervention website. Participants received an email prompt after 2, 4, or 6 weeks. In addition to these 6 experimental conditions, a control condition was included, no prompt (NP). People allocated to the control condition did not receive an additional prompt and were only encouraged at baseline to re-use the program. To participate in the program, participants had to register by using a personal log-in code and password. Immediately after registration for the program, participants were randomly allocated to 1 of the 7 study conditions. Randomization occurred at the respondent level by means of a computer software randomization device. Data for the present study were collected from March to July 2011.

Email Prompt

People who were allocated to the experimental conditions of the study received an email prompting them to re-use the program. Periodic prompts should be distinguished from reminders, which are also often used in Internet-delivered interventions. Sending reminders is a technique that is used to increase response rates in online data collection by proactively stimulating participation among nonresponders and is used to prevent drop-out attrition [21,43]. Periodic prompts, on the other hand, are used to boost re-use of the intervention content by approaching all participants and are used to prevent non-usage attrition [21]. All email prompts used in the present study contained standard content. This standard email opened with a personalized greeting and reminded people about their first visit to the program. Subsequently, people were invited to re-use the program to obtain information about their current health status and to monitor their progress. Participants were also given the opportunity to receive additional iterative health advice on the health behavior(s) selected at baseline or on a new behavior. Finally, to facilitate logging in to the program, the email also contained details about their personal log-in information (username and password). The email concluded with greetings from the research team and contact information. Half of all people in the experimental conditions received an email that also contained additional content (SP+). This additional content consisted of a preview of new information that was placed on the program website since they last visited it. This information referred to nutrition and provided examples of healthy food alternatives that were available for that current season (eg, spring/summer).

Internet-Delivered Computer-Tailored Lifestyle Program

The CT program integrated established CT programs tested and proven to be effective in randomized controlled trials for increasing smoking cessation, promoting the intake of fruit and vegetables, increasing the level of physical activity, and reducing the consumption of alcohol [44-48]. The program used a dual approach to guide people toward behavior change. First, awareness of participants' current health behavior status was increased by comparing their status to the Dutch public health guidelines set for these health behaviors, such as being moderately physically active for 30 minutes at least 5 days a week, eating 2 pieces of fruit per day, eating 200 grams of vegetables per day, not drinking more than 1 (women) or 2 (men) glasses of alcohol a day, and not smoking. Second, assistance was provided in changing participants' health behavior by using CT modules available per behavior. The modules used a fixed, gradual approach consisting of 4 steps, guiding people toward behavior change based on the Integrated Model for exploring motivational and behavioral change (I-Change Model) [49]. Focus was on pros and cons of the desired behavior (step 1), the role of significant persons in the direct environment (step 2), preparatory plans assisting people to start changing their behavior (step 3), and coping plans to help them overcome difficult situations and prevent relapse (step 4). Within the modules, all health advice was adapted to the individual's characteristics by considering demographic, behavioral, and cognitive characteristics [50-52]. The Internet-delivered CT lifestyle program is described in more detail elsewhere [41,42].

The program was embedded in a website designed for the current study. This website entailed general information considering a healthy lifestyle and the selected health behaviors. Furthermore, the website provided specific information regarding the project, contained a direct link to the CT program, and provided background information regarding the study and the research team. During the study, new information (eg, advice, supporting messages, recipes, and facts) was added to the website.

Measures

Participants in the experimental conditions (SP and SP+) received an email prompting them to re-use the intervention. Participants had to take 2 steps to re-use the program in order to self-monitor their level of behavior change and obtain additional CT advice. The first was clicking on the URL provided in the email prompt. Clicking on the URL was tracked by referral ID codes that were integrated in the URL. The second step was logging in to the CT program after arriving on the program website. Log-ins were objectively monitored during a 2- week period and log-in dates were compared to baseline dates to determine revisits.

To describe characteristics of program visitors, age, gender, and educational level of participants was assessed (1/low: no education or lower vocational school; 2/medium: secondary vocational school or high school; 3/high: higher professional education or university). Health behavior status consisted of information regarding the 5 key behaviors. Physical activity was measured by the Short Questionnaire to Assess Health-enhancing physical activity (SQUASH) [53] and guideline adherence was calculated following procedures used by Ainsworth et al (2000) [54]. Fruit consumption was measured by using a 4-item Food Frequency Questionnaire (FFQ) assessing weekly amount of fruit and fruit juice intake [55], whereas vegetable consumption was measured using a 4-item FFQ assessing the weekly amount of consumed boiled or baked vegetables as well as salad or raw vegetables [55]. The consumption of alcohol was measured by the Dutch Quantity-Frequency-Variability (QFV) questionnaire [56]. Finally, smoking status was assessed by asking participants whether they smoked, what they smoked (cigarettes, cigars, pipe tobacco), and how much they smoked per day (cigarettes) and per week (cigars/packets pipe tobacco) [57]. For each health behavior, a new variable was created to indicate whether participants met the Dutch guidelines provided for these behaviors (0=no; 1=yes).

Statistical Analysis

General descriptive statistics were calculated to describe characteristics of program visitors, as well as main findings concerning adherence to the public health guidelines. Baseline differences between the intervention groups were calculated using the Chi-square test for dichotomous and categorical variables and 1-way analysis of variance (ANOVA) for continuous variables. Finally, logistic regression analyses were conducted. There were 2 dependent variables: whether participants clicked on the URL (0=no; 1=yes) and whether participants logged in to the CT program (0=no; 1=yes). Prompt content (dummy coded, SP=0; SP+=1) and prompt timing (dummy coded with 2 weeks as the reference category and 4 weeks as the reference category) and the interactions between these variables were used as predictors in the initial model of each dependent variable. Furthermore, because women and people with a higher educational level and age are more likely to use Internet-delivered lifestyle interventions [17,58], age, gender, and educational level were included in the models as possible covariates. An alpha of .05 was used to indicate statistical significance. All statistical analyses were done with the program SPSS 17.0 (SPSS Inc, Chicago, IL, USA).

Results

In total, 240 participants visited the program, of which 73.3% (176/240) were female. Participants were randomly allocated to each of the 7 study conditions and a randomization check revealed that females were equally distributed (χ^2_6= 9.1, *P*=.17). Furthermore, there were no significant differences between the groups regarding educational level (χ^2_{12}= 9.3, *P*=.68) and age ($F_{6,233}$= 0.464, *P*=.84). Overall visitors had a mean age of 50 years (SD 14.99) and most were medium (100/240, 41.7%) to highly educated (111/240, 46.3%). Regarding the 5 health behaviors included in the program, 12.9% (31/240) did not comply with the Dutch guidelines of at least 30 minutes of moderately intensive physical activity on at least 5 days of the week. With regard to fruit and vegetable intake, 50.8% (122/240) and 60.0% (144/240) did not adhere to the Dutch guidelines of at least 2 pieces of fruit and at least 200 gram of vegetables each day, respectively. Approximately 1 out of 10 participants indicated that they smoked (10.0%, 24/240), and 22.1% (53/240) did not comply with the Dutch guidelines for alcohol intake.

A total of 206 participants received an email prompting them to re-use the program. Of this sample, 53 participants (25.7%) reacted to this email by clicking on the URL (step 1), whereas 25 participants (12.1%) actually logged in to the program (step 2). All results concerning clicking on the URL and logging in to the program are described per study group in Table 1.

Table 1. Participants who clicked on the URL (step 1) and logged in to the program (step 2) per study condition (N=240).

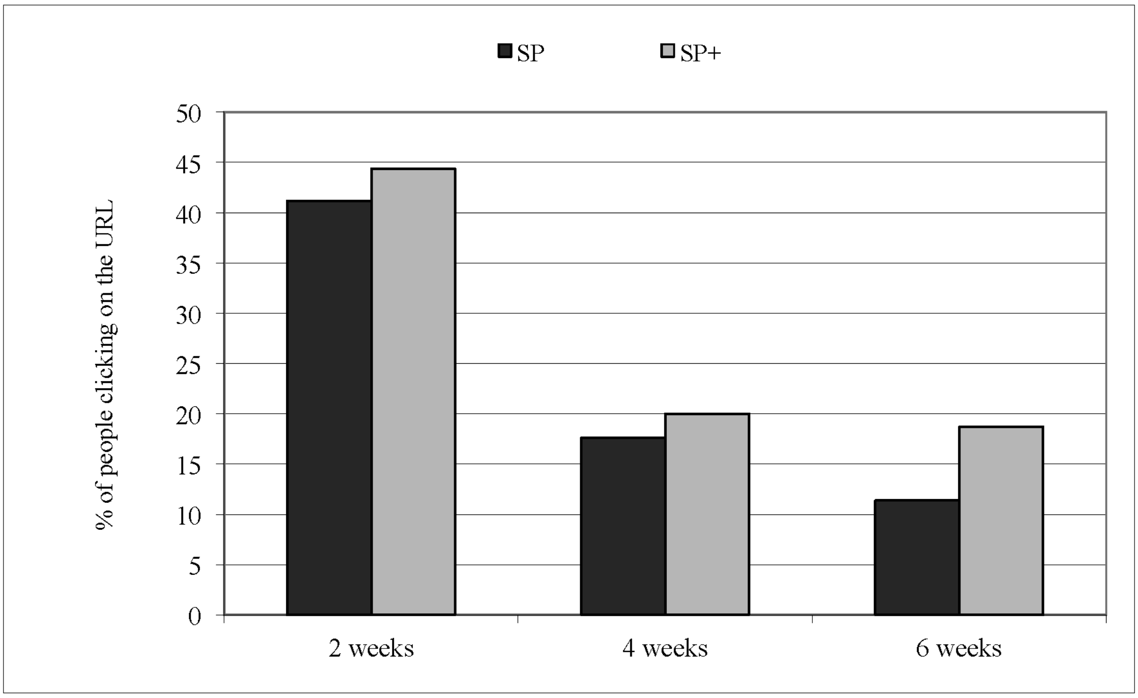
Dependent variable	Condition, n (%)						
	Standard content			Standard+ content			No prompt
	2 weeks n=34	4 weeks n=34	6 weeks n=35	2 weeks n=36	4 weeks n=35	6 weeks n=32	n=34
Click on the URL	14 (41.2)	6 (17.6)	4 (11.4)	16 (44.4)	7 (20.0)	6 (18.8)	—
Log in to the program	6 (17.6)	1 (2.9)	1 (2.9)	10 (27.8)	7 (20.0)	0 (0.0)	2 (5.9)

Step 1: Clicking on the URL

There was no significant interaction between prompt content and prompt timing (Figure 1) with regard to clicking on the URL. Therefore, interaction terms were excluded from the remaining models and only main effects are reported.

Analyses of main effects indicated that there was a significant effect of prompt timing (χ^2_2 = 15.2, *P* = <.001). Participants who received an email prompt 2 weeks after their first visit, clicked on the URL significantly more often compared with participants that received the prompt after 4 weeks (odds ratio [OR] 3.069, 95% CI 1.392-6.765, *P*=.005) and after 6 weeks (OR 4.471, 95% CI 1.909-10.471, *P*=.001). There was no significant difference in reaction to the email prompt between participants receiving the prompts after 4 weeks, compared with participants receiving the prompts after 6 weeks. Also, no main effects of prompt content could be detected (see Table 1).

Figure 1. Percentage of participants given standard prompts (SP) and standard+ prompts (SP+) who clicked on the URL at different levels of prompt timing.



Logging In to the Program

With regard to logging in to the program, no significant interaction between prompt content and prompt timing was found (Figure 2). Therefore, interaction terms were again excluded from the remaining models and only main effects are reported.

With regard to logging in to the program, analyses of main effects indicated that there was a significant effect of prompt timing (χ^2_2 = 16.5, *P*<.001). Participants receiving an email prompt 2 weeks after their first visit, logged in to the program significantly more often compared to participants receiving the prompt after 6 weeks (OR 16.356, 95% CI 2.071-129.196, *P*=.008). There was, however, no significant difference between participants receiving the prompts after 4 weeks compared with participants receiving the prompts after 2 or 6 weeks. With regard to prompt content, a trend was observed. Participants receiving a prompt with additional content were more likely to log in to the program compared to participants who received a prompt with standard content. However, this result was not statistically significant (OR 2.286, 95% CI 0.892-5.856, *P*=.09).

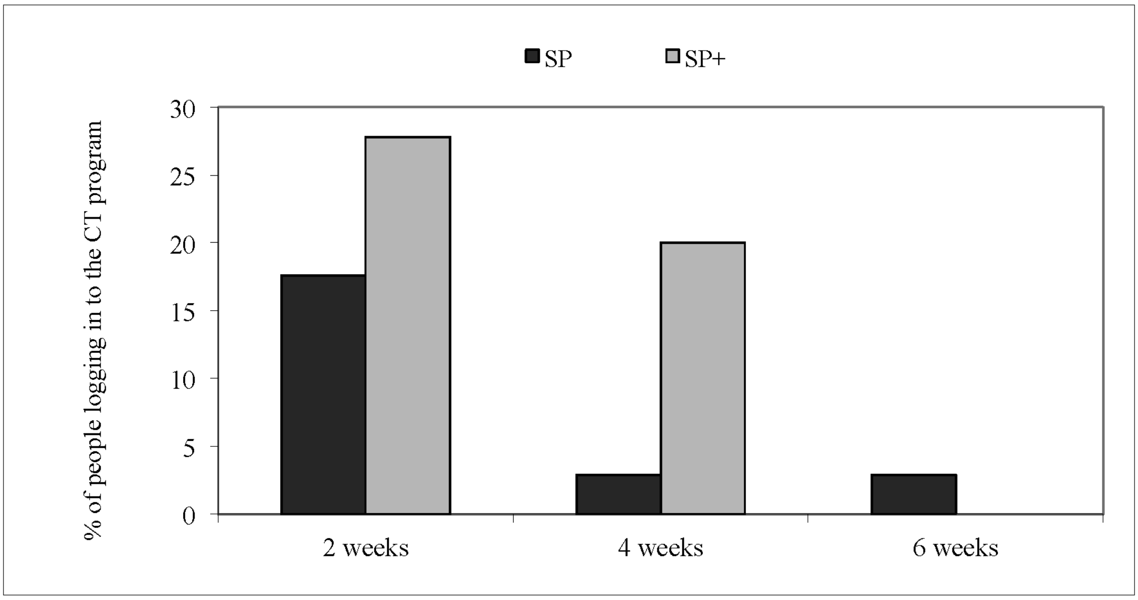
Table 2. Effects of prompt content and timing on clicking on the URL (step 1) and logging in to the program (step 2).

Dependent variable	Condition	OR	95% CI	*P*
Clicking on the URL				
	Prompt content (SP versus SP+)[a]	1.278	0.652-2.505	.48
	Prompt timing (2 week vs 4 week)	3.069	1.392-6.765	.005
	Prompt timing (2 week vs 6 week)	4.471	1.909-10.471	.001
	Prompt timing (4 week vs 6 week)	1.457	0.575-3.689	.43
Logging in to the program				
	Prompt content (SP vs SP+)[a]	2.286	0.892-5.856	.09
	Prompt timing (2 week vs 4 week)	2.144	0.822-5.593	.12
	Prompt timing (2 week vs 6 week)	16.356	2.071-129.196	.008
	Prompt timing (4 week vs 6 week)	0.131	0.016-1.096	.06

[a] SP: standard prompt content; SP+: standard prompt content complemented with a preview of new website content.

Finally, all experimental groups were compared with the control condition (NP) that did not receive an additional email prompt to boost revisits to the intervention. Results from this analysis revealed that participants who received an email prompt containing additional content after 2 weeks were significantly more likely to log in to the program (OR 6.059, 95% CI 1.195-30.726, *P*<.001).

Figure 2. Percentage of participants given standard prompts (SP) and standard+ prompts (SP+) who logged in to the program at different levels of prompt timing.



Discussion

This study aimed at maximizing the potential of email prompts by focusing on prompt content and timing. A single email prompt was used in the context of an Internet-delivered CT lifestyle program and was aimed at stimulating re-use to increase the dose of the intervention and prevent non-usage attrition. Firstly, results indicated that sending a prompt 2 weeks after the first visit was more effective compared with using a longer time period. This effect not only referred to clicking on the URL (step 1), but also to the proportion of log-ins to the program (step 2). Secondly, a trend was observed indicating that adding a preview of new website content to a standard prompt increased its effectiveness in persuading people to log in to the program. Finally, sending a prompt with additional content after a 2-week period significantly increased program log-ins compared to using a reactive approach in which no additional prompts were used.

The effectiveness of using a relatively brief interval for sending a prompt may indicate that people tend to forget about the program. Therefore, the interval between engaging in a first visit and receiving an email prompt to boost re-use should be kept relatively short. Since the costs of sending a large number of email messages are relatively low [59], this strategy allows for sending several emails at relatively short intervals. However, when using an email prompt to boost re-use of the program, one must keep the goal of the revisit in mind. Within the current program, the prompt was used to remind people about their previous visit to the program and to persuade them to re-use the program to monitor their progress and obtain additional feedback. Although results from this study suggest that prompt timing should be short, one must allow participants enough time to actually follow the obtained advice and develop strategies to positively change their lifestyle [32].

Results from the current study also show a trend toward the efficacy of adding additional content to a standard prompt. However, this trend was only present for logging in to the program (step 2) and not for clicking on the URL (step 1). A possible explanation for this effect on log-ins might be the repeated exposure to new website content. People who received an email prompt containing a preview of new website content were exposed to this website content twice: within the email prompt and on the program website. According to ease of processing theory [60], repeated exposure to information allows for more fluent processing of the information. This extended degree of exposure might have enhanced the ease of processing the new information added to the website. Because this new information was related to health and emphasized the importance of having a healthy lifestyle, this information could have persuaded people to monitor their current lifestyle behavior and obtain additional advice about how to positively change their behavior.

With regard to clicking on the URL (step 1), significant effects of adding a peripheral cue to a standard prompt failed to occur. People were significantly more likely to respond to emails that were sent at a relatively short interval, irrespective of the email content. A possible explanation for this finding might be that participants consented to participate in the present study after being informed about the possibility of receiving an additional email to boost program use. Within the field of e-marketing, the term *opt-in email* is used when receivers have agreed to the receipt of an email [34,61]. Research has indicated that opt-in emails are largely accepted among receivers and are proven to be more effective compared to spam email (receiver has not agreed to receiving an email) when used to persuade people to undertake action (eg, clicking on a URL to re-use the program) [34,61]. Therefore, knowledge about receiving an email prompt might have increased acceptance of the email among participants and might have resulted in the desired behavior (clicking on the URL) irrespective of the email content.

Strengths, Limitations, and Recommendations

The current study primarily focused on the effectiveness of using a single email prompt with regard to clicking on the URL and logging in to the program. Combined with results from our previous study [17], important information regarding the overall effectiveness of using a single email prompt as well as the optimal prompt content and timing was obtained. To our knowledge, this is one of the first studies that systematically studied optimal content and timing of a single email prompt used to boost re-use of a multiple CT lifestyle program offered through the Internet. Although these results suggest that sending a single email prompt is an effective strategy to boost revisits to interventions using multiple computer-tailoring, this is only a first step, and the findings should be interpreted keeping several limitations in mind. Firstly, the current study focused solely on sending an email prompt on 1 occasion, namely after 2, 4, or 6 weeks. However, achieving health behavior change is a complex and enduring process that requires continuous guidance [22,26]. Repeated visits to interventions using multiple computer-tailoring are imperative to ensure notification of the entire intervention content and involvement in its effective components [27,28] and must, therefore, be encouraged. Additional research should go one step further and focus on timing intervals and prompt frequency. Prompting for revisits shortly after a first visit indeed stimulated revisits to the program. However, repeatedly sending prompts at a certain interval to boost multiple revisits might have different or even reverse effects because responses to multiple prompts tend to gradually decline over time [62]. Secondly, effectiveness of the use of an email prompt largely depends on the degree to which the content is actually read and processed by receivers. Within the current study, no objective measure of the degree to which emails were read was obtained. Therefore, results on prompt content should be interpreted with caution. Future research should put additional effort in assessing the degree to which email content is actually read and processed by participants. Furthermore, additional research is needed to investigate the effects of using periodic prompts on behavior change and to investigate whether these effects remain present over a sustained period of time. Thirdly, the sample size for the present study was relatively small. As a result, the study might be relatively underpowered to study interaction effects and to indicate significant main effects. Furthermore, regression analyses resulted in several substantial confidence intervals. Although these results should be interpreted with caution, they still provide valuable information regarding the effectiveness of timing and content of a single email prompt as a strategy to boost program revisits. Fourthly, a related limitation is the representativeness of the study sample. This sample does not provide a good cross-section of the general Dutch population [63] because participants were mainly female, middle-aged, and medium to highly educated. This may limit possibilities for generalizing our results to the general population. However, previous studies repeatedly indicated that Internet-delivered CT programs tend to predominantly reach women and people who are older and higher educated [17,58,64-67]. As a consequence, the current sample corresponds to the subgroup of people known to be reached by Internet-delivered CT programs, although it does not represent the general population. Therefore, the obtained results are valuable in the context of CT lifestyle programs that are offered through the Internet.

Conclusions

We found that using relatively short prompt timing (2 weeks) resulted in more positive effects compared to using a longer time period. This effect not only referred to clicking on the URL, but also to the proportion of log-ins to the program. Although a trend was observed about the effectiveness of adding a preview of new website content to a standard prompt, this only referred to persuading people to log in to the program. The findings of this study underline the importance of sending an email prompt relatively shortly after a first visit to the program. Furthermore, it is important to further focus on prompt content because the addition of a peripheral cue may add to the prompt's effectiveness.

Acknowledgments

This study was funded by ZonMw, the Netherlands Organisation for Health Research and Development (grant number: 120710004). The intervention was developed and implemented by the Department of Health Promotion of CAPHRI, Maastricht University. The authors would like to thank the Regional Public Health Services of the Dutch province of Limburg (GGD Limburg Noord) for the close collaboration in recruiting the participants and collecting the data. Furthermore, we would like to thank all members of the Health Communication meeting for their valuable and constructive comments on an earlier version of this manuscript.

Conflicts of Interest

Hein de Vries is scientific director of Vision2Health, a company that licenses innovative evidence-based computer-tailored health communication tools.

References

1. Brug J, Oenema A, Kroeze W, Raat H. The internet and nutrition education: challenges and opportunities. Eur J Clin Nutr 2005 Aug;59 Suppl 1:S130-7; discussion S138. [doi: 10.1038/sj.ejcn.1602186] [Medline: 16052181]
2. Ritterband LM, Gonder-Frederick LA, Cox DJ, Clifton AD, West RW, Borowitz SM. Internet interventions: in review, in use and into the future. Professional Psychology: Research and Practice 2003;34(5):527-534. [doi: 10.1037/0735-7028.34.5.527]
3. Krebs P, Prochaska JO, Rossi JS. A meta-analysis of computer-tailored interventions for health behavior change. Prev Med 2010;51(3-4):214-221 [FREE Full text] [doi: 10.1016/j.ypmed.2010.06.004] [Medline: 20558196]
4. 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]
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. Neville LM, O'Hara B, Milat AJ. Computer-tailored dietary behaviour change interventions: a systematic review. Health Educ Res 2009 Aug;24(4):699-720 [FREE Full text] [doi: 10.1093/her/cyp006] [Medline: 19286893]
7. Brug J, Glanz K, Van Assema P, Kok G, van Breukelen GJ. The impact of computer-tailored feedback and iterative feedback on fat, fruit, and vegetable intake. Health Educ Behav 1998 Aug;25(4):517-531. [Medline: 9690108]
8. Borland R, Balmford J, Hunt D. The effectiveness of personally tailored computer-generated advice letters for smoking cessation. Addiction 2004 Mar;99(3):369-377. [doi: 10.1111/j.1360-0443.2003.00623.x] [Medline: 14982550]
9. Strecher VJ. Computer-tailored smoking cessation materials: a review and discussion. Patient Educ Couns 1999 Feb;36(2):107-117. [Medline: 10223016]
10. Myung SK, McDonnell DD, Kazinets G, Seo HG, Moskowitz JM. Effects of Web- and computer-based smoking cessation programs: meta-analysis of randomized controlled trials. Arch Intern Med 2009 May 25;169(10):929-937. [doi: 10.1001/archinternmed.2009.109] [Medline: 19468084]
11. Dijkstra A, De Vries H, Roijackers J. Long-term effectiveness of computer-generated tailored feedback in smoking cessation. Health Educ Res 1998 Jun;13(2):207-214 [FREE Full text] [Medline: 10181019]
12. Riper H, Kramer J, Smit F, Conijn B, Schippers G, Cuijpers P. Web-based self-help for problem drinkers: a pragmatic randomized trial. Addiction 2008 Feb;103(2):218-227. [doi: 10.1111/j.1360-0443.2007.02063.x] [Medline: 18199300]
13. White A, Kavanagh D, Stallman H, Klein B, Kay-Lambkin F, Proudfoot J, et al. Online alcohol interventions: a systematic review. J Med Internet Res 2010;12(5):e62 [FREE Full text] [doi: 10.2196/jmir.1479] [Medline: 21169175]
14. Dijkstra A, De Vries H, Roijackers J, van Breukelen G. Tailoring information to enhance quitting in smokers with low motivation to quit: three basic efficacy questions. Health Psychol 1998 Nov;17(6):513-519. [Medline: 9848801]
15. 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]
16. Crutzen R. Hard to Get, Hard to Keep. Dissemination of and Exposure to Internet-Delivered Health Behavior Change Interventions Aimed at Adolescents. Maastricht, the Netherlands: Maastricht University; 2009.
17. Schneider F, van Osch L, Schulz DN, Kremers SP, de Vries H. The influence of user characteristics and a periodic email prompt on exposure to an internet-delivered computer-tailored lifestyle program. J Med Internet Res 2012;14(2):e40 [FREE Full text] [doi: 10.2196/jmir.1939] [Medline: 22382037]
18. 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]
19. Rogers EM. Diffusion of Innovations. New York: Free Press; 2003.
20. 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]
21. 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]
22. 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]
23. 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]
24. 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]

25. 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]
26. 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]
27. 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]
28. Strecher VJ, McClure J, Alexander G, Chakraborty B, Nair V, Konkel J, et al. The role of engagement in a tailored web-based smoking cessation program: randomized controlled trial. J Med Internet Res 2008;10(5):e36 [FREE Full text] [doi: 10.2196/jmir.1002] [Medline: 18984557]
29. Heesch KC, Mâsse LC, Dunn AL, Frankowski RF, Mullen PD. Does adherence to a lifestyle physical activity intervention predict changes in physical activity? J Behav Med 2003 Aug;26(4):333-348. [Medline: 12921007]
30. Schneider F, van Osch L, de Vries H. Identifying factors for optimal development of health-related websites: a delphi study among experts and potential future users. J Med Internet Res 2012;14(1):e18 [FREE Full text] [doi: 10.2196/jmir.1863] [Medline: 22357411]
31. Prochaska JO, Velicer WF, Fava JL, Rossi JS, Tsoh JY. Evaluating a population-based recruitment approach and a stage-based expert system intervention for smoking cessation. Addict Behav 2001;26(4):583-602. [Medline: 11456079]
32. 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]
33. Fry JP, Neff RA. Periodic prompts and reminders in health promotion and health behavior interventions: systematic review. J Med Internet Res 2009;11(2):e16 [FREE Full text] [doi: 10.2196/jmir.1138] [Medline: 19632970]
34. Raad M, Yeassan NM, Alam GM, Zaidan BB, Zaidan AA. Impact of spam advertisement through e-mail: A study to assess the influence of the anti-spam on the e-mail marketing. African Journal of Business Management 2010;4(11):2362-2367.
35. Merisavo M, Raulas M. The impact of e-mail marketing on brand loyalty. Journal of Product & Brand Management 2004;13(7):498-505. [doi: 10.1108/10610420410568435]
36. Göritz A, Crutzen R. Reminders in web-based data collection: Increasing response at the price of retention. American Journal of Evaluation 2012;33:240-250 [FREE Full text] [doi: 10.1177/1756283X10363751] [Medline: 22973420]
37. Kongsved SM, Basnov M, Holm-Christensen K, Hjollund NH. Response rate and completeness of questionnaires: a randomized study of Internet versus paper-and-pencil versions. J Med Internet Res 2007;9(3):e25 [FREE Full text] [doi: 10.2196/jmir.9.3.e25] [Medline: 17942387]
38. Svensson M, Svensson T, Hansen AW, Trolle Lagerros Y. The effect of reminders in a web-based intervention study. Eur J Epidemiol 2012 May;27(5):333-340. [doi: 10.1007/s10654-012-9687-5] [Medline: 22531973]
39. Petty R, Cacioppo JT. The elaboration likelihood of persuasion. Advances in Experimental Social Psychology 1986;19:123-205.
40. Menon S, Soman D. Managing the power of curiosity for effective web advertising strategies. Journal of Advertising 2002;31(3):1-14.
41. Schneider F, van Osch LA, Kremers SP, Schulz DN, van Adrichem MJ, de Vries H. Optimizing diffusion of an online computer tailored lifestyle program: a study protocol. BMC Public Health 2011;11:480 [FREE Full text] [doi: 10.1186/1471-2458-11-480] [Medline: 21689412]
42. Schulz DN, Kremers SP, van Osch LA, Schneider F, van Adrichem MJ, de Vries H. Testing a Dutch web-based tailored lifestyle programme among adults: a study protocol. BMC Public Health 2011;11:108 [FREE Full text] [doi: 10.1186/1471-2458-11-108] [Medline: 21324181]
43. Göritz AS, Crutzen R. Reminders in web-based data collection: increasing response at the price of retention. American Journal of Evaluation 2012;33:240-240 [FREE Full text]
44. Smeets T, Kremers SP, Brug J, de Vries H. Effects of tailored feedback on multiple health behaviors. Ann Behav Med 2007 Apr;33(2):117-123. [doi: 10.1080/08836610701307801] [Medline: 17447863]
45. Brug J, Steenhuis I, van Assema P, Glanz K, De Vries H. Computer-tailored nutrition education: differences between two interventions. Health Educ Res 1999 Apr;14(2):249-256 [FREE Full text] [Medline: 10387504]
46. 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]
47. Smit ES, de Vries H, Hoving C. The PAS study: a randomized controlled trial evaluating the effectiveness of a web-based multiple tailored smoking cessation programme and tailored counselling by practice nurses. Contemp Clin Trials 2010 May;31(3):251-258. [doi: 10.1016/j.cct.2010.03.001] [Medline: 20215047]
48. van Keulen HM, Mesters I, Brug J, Ausems M, Campbell M, Resnicow K, et al. Vitalum study design: RCT evaluating the efficacy of tailored print communication and telephone motivational interviewing on multiple health behaviors. BMC Public Health 2008;8:216 [FREE Full text] [doi: 10.1186/1471-2458-8-216] [Medline: 18565222]

49. de Vries H, Mudde A, Leijs I, Charlton A, Vartiainen E, Buijs G, et al. The European Smoking Prevention Framework Approach (EFSA): an example of integral prevention. Health Educ Res 2003 Oct;18(5):611-626 [FREE Full text] [Medline: 14572020]
50. de Vries H, Brug J. Computer-tailored interventions motivating people to adopt health promoting behaviours: introduction to a new approach. Patient Educ Couns 1999 Feb;36(2):99-105. [Medline: 10223015]
51. Dijkstra A, De Vries H. The development of computer-generated tailored interventions. Patient Educ Couns 1999 Feb;36(2):193-203. [Medline: 10223023]
52. Kreuter MW, Farrell DW, Olevitch LR, Brennan LK. Tailoring Health Messages: Customizing Communication with Computer Technology. Mahwah, NJ: Lawrence Erlbaum Associates; 2000.
53. Wendel-Vos GC, Schuit AJ, Saris WH, 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]
54. Ainsworth BE, Haskell WL, Whitt MC, Irwin ML, Swartz AM, Strath SJ, et al. Compendium of physical activities: an update of activity codes and MET intensities. Med Sci Sports Exerc 2000 Sep;32(9 Suppl):S498-S504. [Medline: 10993420]
55. Lokale en nationale Monitor Gezondheid [Local and national Health Monitor]. 2005. Indicatoren voor de Monitor Volksgezondheid [Indicators of the Public Health Monitor] URL: https://www.monitorgezondheid.nl/volksindicatoren.aspx [accessed 2013-01-23] [WebCite Cache ID 6DtQ6bWPt]
56. Lemmens P, Tan ES, Knibbe RA. Measuring quantity and frequency of drinking in a general population survey: a comparison of five indices. J Stud Alcohol 1992 Sep;53(5):476-486. [Medline: 1405641]
57. de Vries H, Kremers SP, Smeets T, Brug J, Eijmael K. The effectiveness of tailored feedback and action plans in an intervention addressing multiple health behaviors. Am J Health Promot 2008;22(6):417-425. [doi: 10.4278/ajhp.22.6.417] [Medline: 18677882]
58. 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]
59. Merisavo M, Raulas M. The impact of e-mail marketing on brand loyalty. Journal of Product & Brand Management 2004;13(7):498-505. [doi: 10.1108/10610420410568435]
60. Whittlesea BWA, Jacoby JL, Girardi K. Illusions of immediate memory: evidence of an attributional basis for feelings of familiarity and perceptual quality. Journal of Memory and Language 1990;29:716-732. [doi: 10.1016/0749-596X(90)90045-2]
61. Barnes SJ, Scornavacca E. Mobile marketing: the role of permission and acceptance. International Journal of Mobile Communications 2004;2(2):128-139. [doi: 10.1504/IJMC.2004.004663]
62. 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]
63. Centraal Bureau voor de Statistiek (CBS) [Statistics Netherlands]. 2012. URL: http://www.cbs.nl/nl-NL/menu/themas/bevolking/nieuws/default.htm [accessed 2013-01-22] [WebCite Cache ID 6DrhPqFDm]
64. Spittaels H, De Bourdeaudhuij I, Brug J, Vandelanotte C. Effectiveness of an online computer-tailored physical activity intervention in a real-life setting. Health Educ Res 2007 Jun;22(3):385-396 [FREE Full text] [doi: 10.1093/her/cyl096] [Medline: 16971674]
65. Vandelanotte C, De Bourdeaudhuij I, Sallis JF, Spittaels H, Brug J. Efficacy of sequential or simultaneous interactive computer-tailored interventions for increasing physical activity and decreasing fat intake. Ann Behav Med 2005 Apr;29(2):138-146. [doi: 10.1207/s15324796abm2902_8] [Medline: 15823787]
66. Spittaels H, De Bourdeaudhuij I. Who participates in a computer-tailored physical activity program delivered through the Internet? A comparison of participants' and non-participants' characteristics. Int J Behav Nutr Phys Act 2007;4:39 [FREE Full text] [doi: 10.1186/1479-5868-4-39] [Medline: 17880696]
67. Schulz DN, Schneider F, de Vries H, van Osch LA, van Nierop PW, Kremers SP. Program completion of a web-based tailored lifestyle intervention for adults: differences between a sequential and a simultaneous approach. J Med Internet Res 2012;14(2):e26 [FREE Full text] [doi: 10.2196/jmir.1968] [Medline: 22403770]

Abbreviations

ANOVA: analysis of variance
CT: computer tailored
ELM: elaboration likelihood model
FFQ: Food Frequency Questionnaire
NP: no prompt
OR: odds ratio
QFV: Quantity-Frequency-Variability
RPHS: Regional Public Health Service
SP: standard prompt

SP+: standard+ prompt
SQUASH: Short Questionnaire to Assess Health-enhancing physical activity
URL: uniform resource locator

Edited by G Eysenbach; submitted 30.04.12; peer-reviewed by ML Lustria; comments to author 27.06.12; revised version received 30.08.12; accepted 05.10.12; published 31.01.13.

Please cite as:
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
URL: http://www.jmir.org/2013/1/e23/
doi:10.2196/jmir.2151
PMID:23363466

Original Paper

Mobile Phone Applications for the Care and Prevention of HIV and Other Sexually Transmitted Diseases: A Review

Kathryn E Muessig[1], PhD; Emily C Pike[1], BS; Sara LeGrand[2], PhD; Lisa B Hightow-Weidman[1], MD, MPH

[1]Department of Medicine, The University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[2]Center for Health Policy and Inequalities Research, Duke University, Durham, NC, United States

Corresponding Author:
Kathryn E Muessig, PhD
Department of Medicine
The University of North Carolina at Chapel Hill
Bioinformatics Bldg, CB #7030
130 Mason Farm Rd
Chapel Hill, NC, 27599
United States
Phone: 1 443 320 3152
Fax: 1 919 966 6714
Email: kmuessig@med.unc.edu

Abstract

Background: Mobile phone applications (apps) provide a new platform for delivering tailored human immunodeficiency virus (HIV) and sexually transmitted disease (STD) prevention and care.

Objective: To identify and evaluate currently available mobile phone apps related to the prevention and care of HIV and other STDs.

Methods: We searched the Apple iTunes and Android Google Play stores for HIV/STD-related apps, excluding apps that exclusively targeted industry, providers, and researchers. Each eligible app was downloaded, tested, and assessed for user ratings and functionality as well as 6 broad content areas of HIV prevention and care: HIV/STD disease knowledge, risk reduction/safer sex, condom promotion, HIV/STD testing information, resources for HIV-positive persons, and focus on key populations.

Results: Search queries up to May 2012 identified 1937 apps. Of these, 55 unique apps met the inclusion criteria (12 for Android, 29 for iPhone, and 14 for both platforms). Among these apps, 71% provided disease information about HIV/STDs, 36% provided HIV/STD testing information or resources, 29% included information about condom use or assistance locating condoms, and 24% promoted safer sex. Only 6 apps (11%) covered all 4 of these prevention areas. Eight apps (15%) provided tools or resources specifically for HIV/STD positive persons. Ten apps included information for a range of sexual orientations, 9 apps appeared to be designed for racially/ethnically diverse audiences, and 15 apps featured interactive components. Apps were infrequently downloaded (median 100-500 downloads) and not highly rated (average customer rating 3.7 out of 5 stars).

Conclusions: Most available HIV/STD apps have failed to attract user attention and positive reviews. Public health practitioners should work with app developers to incorporate elements of evidence-based interventions for risk reduction and improve app inclusiveness and interactivity.

(J Med Internet Res 2013;15(1):e1) doi:10.2196/jmir.2301

KEYWORDS

HIV; technology; mobile phone applications

Introduction

Mobile phone health interventions are increasingly being used for the prevention and care of human immunodeficiency virus (HIV) and other sexually transmitted diseases (STDs) [1-3]. These initiatives have been designed to promote prevention messages [4], facilitate test result notification [5,6], improve HIV medication adherence, and increase adherence to clinic appointments [7-12]. Although phone-based interventions have typically used the voice or text-based Short Message Service (SMS) features of mobile phones [1,3,13], the increasing popularity of smartphones and smartphone applications (apps)

[14,15] has greatly expanded the possibilities for phone-based HIV/STD interventions [13]. These interventions are critical for reversing the HIV epidemic; 34.2 million people worldwide are living with HIV and 2.5 million people became newly infected in 2011 alone [16].

Smartphones have revolutionized mobile communication markets by offering enhanced mobile phones featuring improved Internet access and the capacity to perform more advanced computer functions. Consumer research from June 2012 estimates that 54.9% of mobile phone subscribers in the United States own a smartphone, with phones running on Apple's iPhone operating systems (iOS) and Android's operating systems (Android OS) representing over 86% of this market [15]. Smartphone apps are downloadable programs that run on the smartphone's OS, which may include Web-based features. As of March 2012, Apple's iTunes apps had received over 25 billion downloads [14] and the Android Google Play Store exceeded 15 billion app downloads in May 2012 [17]. These numbers continue to rise at over 1 billion app downloads per month for each store [17].

Many of these apps are health related. A search on August 1, 2012, for apps categorized as "health & fitness" and "medical" yielded 13,479 apps available for iPhone consumers [18] and 15,891 apps available for Android consumers [19]. This widespread use of apps provides a promising new platform for delivering tailored HIV prevention messages and interactive care services. But are public health practitioners utilizing this new opportunity? In this paper, we review the characteristics and content of HIV- and STD-related apps that are available through the two primary online app providers: the Apple iTunes Store and the Android Google Play Store. This review describes the current landscape and content of HIV/STD apps, assesses utilization and acceptability of these apps, and provides recommendations to guide the design and development of future apps.

Methods

Search and Screening Strategy

On May 1 to 3, 2012, the following terms were used to search the Apple iTunes Store and the Android Google Play Store: HIV, human immunodeficiency virus, acquired immune deficiency syndrome, sexually transmitted diseases, STD, sexually transmitted infections, STI, sexual health, safe sex, and condom. An app was excluded if it did not include HIV/STD content; exclusively targeted industry, health care, research, or medical professionals; or was not available in English. Two researchers (KEM and ECP) searched each store and compiled lists of all identified apps. App titles and descriptions were screened for relevance and lists of apps to download were created. Lists were compared and differences resolved by a third member of the study team (LBH). Any eligible app that was identified in only one store was searched for by name in the other store to confirm its exclusive availability.

Data Extraction and App Assessment

The following data was extracted for each eligible app on May 1 to 3, 2012: name, platform (iPhone, Android, or both), category as defined by site (eg, medical, education, health and fitness, lifestyle), description of app content as provided by site, price, user star rating, number of customer downloads (available for Android only), number of customer ratings (available for iPhone only), and the date the app was last updated by the developer. For apps that were available in both the Google Play and iTunes stores, we recorded separate price and user rating information for each store. From May 7 to 24, 2012, we downloaded each eligible app (both free and fee-based) and tested all features and functions. If an app had both a free and fee-based version, only the free version was downloaded and assessed.

To evaluate the apps we identified 6 broad content areas for advancing the prevention and treatment of HIV based on the World Health Organization's Global Health Sector Strategy on HIV/AIDS 2011-2015 [20] and the National AIDS Strategy for the United States [21]. These content areas include HIV/STD knowledge and awareness, behavior change/risk reduction/safer sex promotion, condom use/promotion, HIV/STD testing, resources and linkage to care for HIV/STD-positive persons, and intensified or focused efforts for key communities/populations. Each app was assessed for inclusion of the 5 content areas, operationalized as:

1. Does the app provide information about HIV or other STDs?

2. Does the app provide information or descriptions about ways to reduce the risk of sexually transmitting or acquiring HIV/STDs?

3. Does the app provide information about how to use or obtain male or female condoms?

4. Does the app promote or provide information or resources about HIV/STD testing?

5. Does the app provide resources specifically for HIV/STD-positive persons? (Features and resources provided through apps for HIV/STD-positive persons were tested and described.)

If the app also covered other content areas, these were listed and described. These descriptions were then compiled and grouped, resulting in the following additional 5 content categories: drug and alcohol risk, relationships, HIV/STD news, HIV/STD stigma, and an HIV/STD status verification service.

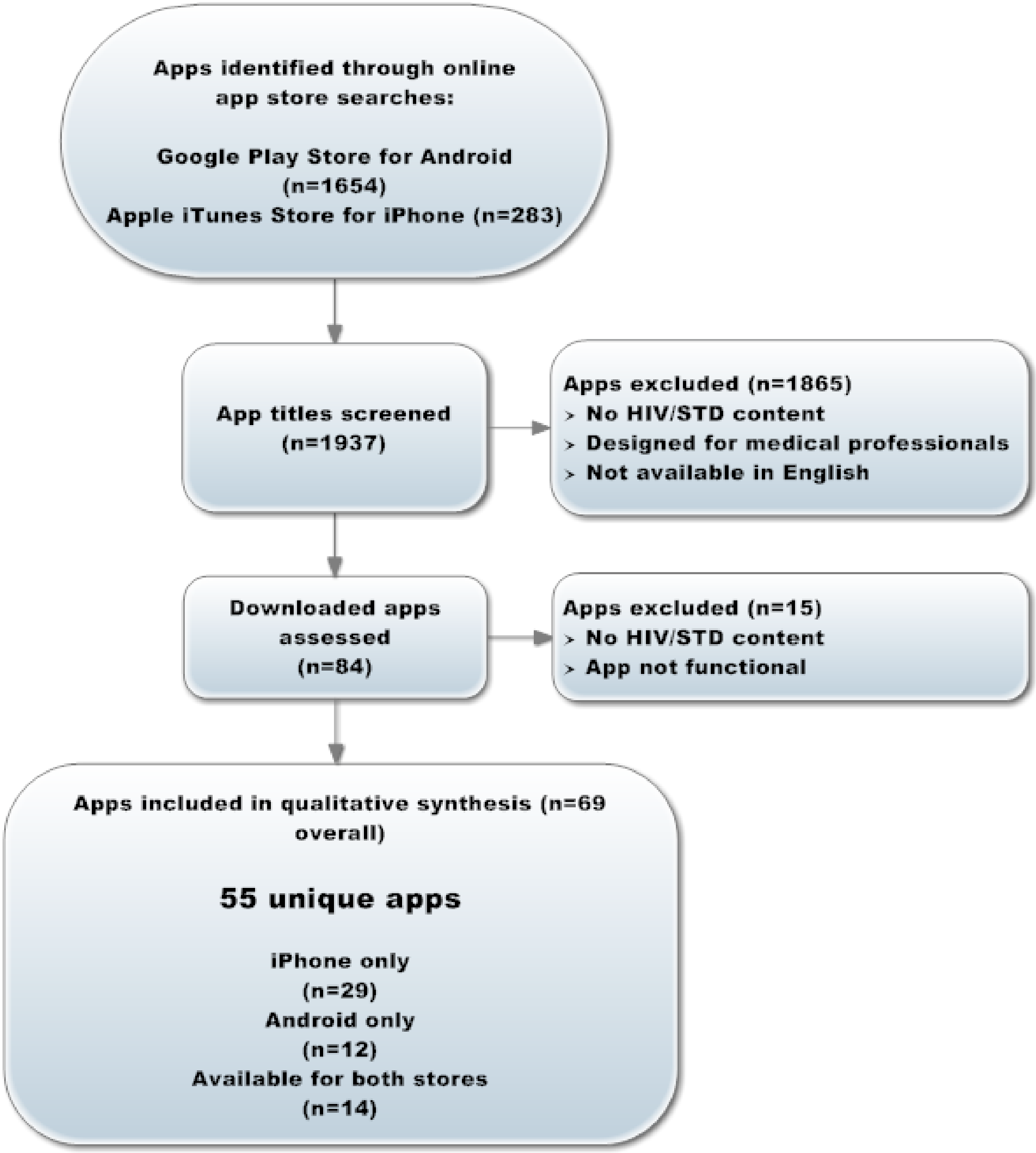
To assess whether the app fulfilled the criteria of intensified or focused efforts for key communities or populations, we reviewed the text and images used in each app. Key populations for HIV prevention vary depending on the local epidemic context. Thus, this assessment included an open-ended description of text and images used and a subjective assessment of whether the app appeared to be inclusive or tailored. For example, if an app included images of black men or women among a variety of racial and ethnic images, this app would be characterized as "inclusive." If an app explicitly or exclusively addressed black men or women, this app would be considered "tailored." Finally, we assessed whether each app included any interactive component, such as a game, quiz, diary, or goal tracker.

Results

Search and Screening

Search queries identified 1937 apps. After screening the app titles and the site-provided app descriptions, 84 eligible apps were downloaded for full review (Figure 1). A further 15 apps were then excluded because they did not include HIV/STD-related content or were not functional (list available upon request). A total of 69 apps met the final inclusion criteria: 43 from the Apple iTunes Store and 26 from Android's Google Play Store. Of these, 55 apps were unique with 14 apps available in both stores (29 iTunes only, 12 Google Play only, and 14 available in both stores).

Figure 1. Search and screening process for HIV/STD-related apps.



Descriptive Characteristics

Table 1 provides a summary of app characteristics and Table 2 lists the name, details, and primary content areas of each eligible app. Approximately half (27/55, 49%) of the apps had been updated within the past year (May 2011 to May 2012). Seventy-one percent (39/55) of apps provided information about HIV/STDs and 24% (13/55) provided information about reducing sexual transmission risk. Twenty-nine percent (16/55) of apps included instructions for condom use or assistance locating condoms and 36% (20/55) included information about

HIV/STD testing or resources for finding testing centers. Only 6 apps covered all 4 of these broad areas of HIV/STD prevention (knowledge, risk reduction, condom promotion, and testing): Action for AIDS [22], Pos or Not [23], STD411 [24], STDiQ [25], Safe Sex Tips [26], and Sexual Health Guide [27]. In addition, one app supported a third-party service that provided an electronic certificate verifying negative HIV/STD test results (Chec-Mate) [28]. Five (9%) apps addressed (noninjection) drug or alcohol use and HIV/STD risk. A small number of apps discussed HIV/STDs in the context of relationships (n=5), provided links to online news stories featuring HIV/STDs (n=4), or addressed HIV stigma (n=1).

Table 1. Summary of mobile phone applications (apps) for HIV/STD prevention and care (N=55).

App characteristic		n	%
Phone platform			
	iPhone	29	52.7
	Android	12	21.8
	Both	14	25.5
App price[a] (US $)			
	Free	46	66.7
	0.99	14	20.3
	1.00-9.99	9	13.0
Customer app rating[a]			
	Unrated[b]	31	44.9
	0.0-3.9 stars	16	23.1
	4.0-5.0 stars	22	31.8
Number of app downloads[c]			
	0-100	12	46.2
	101-500	3	11.5
	501-999	1	3.8
	1000-9999	5	19.2
	> 10,000	5	19.2
Updated since May 2011			
	Yes	27	49.1
	No	28	50.9
App category			
	Brain and puzzle	1	1.8
	Casual	1	1.8
	Education	6	10.9
	Entertainment	1	1.8
	Games	1	1.8
	Health and fitness	20	36.4
	Lifestyle	10	18.2
	Medical	13	23.6
	Social networking	1	1.8
	Utilities	1	1.8
Focused for particular populations[d]			
	Geographically focused	13	23.6
	Inclusive of black or Latino users	9	16.4
	Tailored for black or Latino users	2	3.6
	Information for MSM	7	12.7
	Tailored for MSM	2	3.6
HIV/STD-focused content/features[d]			
	HIV/STD information	39	70.9
	HIV/STD testing	20	36.4

App characteristic		n	%
	Condom use, condom locator	16	29.1
	Game/quiz/risk assessment	15	27.3
	Safer sex	13	23.6
	Tools for HIV/STD-positive persons	8	14.5
	Drugs, alcohol (HIV/STD risk)	6	10.9
	Relationships	5	9.1
	HIV/STD news	4	7.3
	Stigma	1	1.8
	HIV/STD status verification service	1	1.8

[a] Calculated based on 69 total apps because Google Play and iTunes occasionally charged different prices for the same app and received different customer ratings.

[b] Apps that have not received any user ratings.

[c] Information on number of app downloads is only available for apps sold in the Android Google Play Store (n=26).

[d] Each app may be categorized with multiple content areas. Percentages were calculated as percent of unique apps (n=55) that include this content area.

HIV: human immunodeficiency virus; STD: sexually transmitted disease; MSM: men who have sex with men.

Table 2. Characteristics and content of mobile phone HIV/STD applications (apps).

App name	Price (US $)	Rating[a]	Downloads[b]	Targeted population[c]	App focus					
					Knowledge	Safer sex	Condoms[d]	Testing[e]	HIV/ STD+[f]	Other
iPhone apps										
Action For AIDS	0	—	—	Singapore	Y	Y	Y	Y		
AIDSinfo HIV/AIDS Glossary	0	4.5	—	Inclusive: Spanish	Y					
aidsmap news	0	—	—	NT	Y					News
amfAR TestingDay	0	—	—	NT	Y			Y		Quiz
Chec-Mate	1.99[g]	—	—	NT				Y		HIV/STD status verification service
Condom Pro	0	—	—	NT			Y			Game
Condom Truth	0	—	—	NT			Y			
Hiv & Aids Guide	0.99	—	—	Anal sex	Y	Y	Y			
HIV and Your Heart	0	5	—	Inclusive: black women	Y				Y	
HIV Study (AIDS disease)	0.99	—	—	NT	Y			Y		
i*********e-info	0	2.5	—	NT	Y					
iCondom	0	5	—	NT			Y			
iCondom Philly	0	4.5	—	Philadelphia			Y			
It’s Your Shout	0	—	—	England				Y		Quiz
LoveSmarts	0.99	—	—	NT	Y	Y				Game
MASTDinfo	0	—	—	Massachusetts; tailored: MSM	Y			Y		
Men’s Sexual Health	4.99	3	—	Inclusive: black men	Y					
My First Time	0	3	—	China/ Singapore	Y			Y		
Pos or Not	0	4	—	Inclusive: racial/ethnic minorities	Y	Y	Y	Y		Stigma, quiz/game
PozTracker	3.99	—	—	NT	Y				Y	
Safe Sex	0	4	—	NT		Y				
Safesex Guide	1.99	—	—	Denmark	Y	Y				Alcohol, relationships, quiz
Sexual Health News Reader	0.99	—	—	NT	Y					News
STD Guide	1.99	3	—	NT	Y			Y		
STD411	0	2	—	San Francisco; tailored: gay, bi, trans	Y	Y	Y	Y		Risk assessment
STDiQ	0	4	—	Tailored: black, Latino	Y	Y	Y	Y		Alcohol/drugs, risk assessment, relationships

App name	Price (US $)	Rating[a]	Downloads[b]	Targeted population[c]	App focus					
					Knowledge	Safer sex	Condoms[d]	Testing[e]	HIV/STD+[f]	Other
Stop AIDS	0	4	—	Anal sex	Y					
TalkPositive	0	—	—	NT					Y	
TKNO	0	—	—	Inclusive: black men				Y		
Android apps										
About HIV	0	4.5	++	NT	Y			Y		News
Chlamydia	0	—	++	NT	Y			Y		
Guide to STDs	5.00	1	+	NT	Y					
HIV RISK Calculator	0	5	+	Inclusive: MSM, anal sex	Y			Y		Risk assessment
Hook-ups: STD Rally	0	3.5	++++	NT	Y					Game
Kenny Condom	0.99	—	+	Tailored: young African American men			Y	Y		
Know Sexual Health	0	2.7	++++	NT	Y					Quiz
Protection-Sex	0.99	—	+	England	Y	Y	Y			
Safer Sex	0.99	—	+	England; inclusive: gay youth, anal sex	Y	Y	Y			
Sex Detective	0.99	1.9	++++	Inclusive: anal sex	Y					Risk assessment
Sex Guide	0.99	—	+	England; inclusive: gay youth	Y	Y				Drugs, alcohol, relationships
STD Risk Calculator Lite	0	5	++	Inclusive: MSM, anal sex						Risk assessment
Available on iPhone and Android										
AfterSex	0.99/0[h]	3.5/3.6	++++	Inclusive: ethnic Minorities	Y			Y		Risk assessment
Big Night Out	0.99	—/—	+	England	Y		Y			Drugs, alcohol
Birmingham AIDS Outreach (BAO)	0	5/5	+	Inclusive: MSM and males of color	Y			Y	Y	
iStayHealthy	0	—/4.7	++	NT					Y	
Patient Treatment Companion	0	—/—	+	NT					Y	
NYC Condom finder	0	5/4	++++	New York City			Y			
PositiveSingles	0	2/2	+++	NT					Y	
Red Ribbon HIV/AIDS Manager	9.99/1.99[i]	—/4.7	+	NT					Y	
Safe Sex Tips	0/1.14[j]	4/—	+	Inclusive: anal sex	Y	Y	Y	Y		Drugs, alcohol, relationships
Sex Facts	0	3.5/4	+++++	Anal sex	Y					
SeX Factor	0	—/2.2	++++	NT	Y					Game

App name	Price (US $)	Rating[a]	Downloads[b]	Targeted population[c]	App focus					
					Knowledge	Safer sex	Condoms[d]	Testing[e]	HIV/STD+[f]	Other
Sexual Health	0	3/4	+++	NT	Y					
Sexual Health Guide	0	—/5	+++	Ireland; inclusive: anal sex	Y	Y	Y	Y		Drugs, alcohol, relationships, news
STD Glossary	0.99	—/4	++	NT	Y					

[a] —: not yet rated. For apps available for both iPhone and Android, ratings presented as iPhone/Android.

[b] Download data was recorded on May 1-3, 2012, and was only available for Android Google Play store; —: not reported; +: 1-100 downloads; ++: 101-1000 downloads; +++: 1001-5000 downloads; ++++: 5001-50,000 downloads; +++++: 50,000-100,0000 downloads.

[c] NT: not inclusive or tailored for specific geographic area or population(s); MSM: men who have sex with men.

[d] Condom use and condom GPS locators.

[e] Testing promotion and clinic locators.

[f] Information, services, and/or care management tools for persons diagnosed with HIV or other sexually transmitted diseases (STDs).

[g] For 1-year membership.

[h] iPhone version $0.99.

[i] iPhone version $9.99.

[j] iPhone version free.

Apps for Persons Diagnosed With HIV or Other STDs

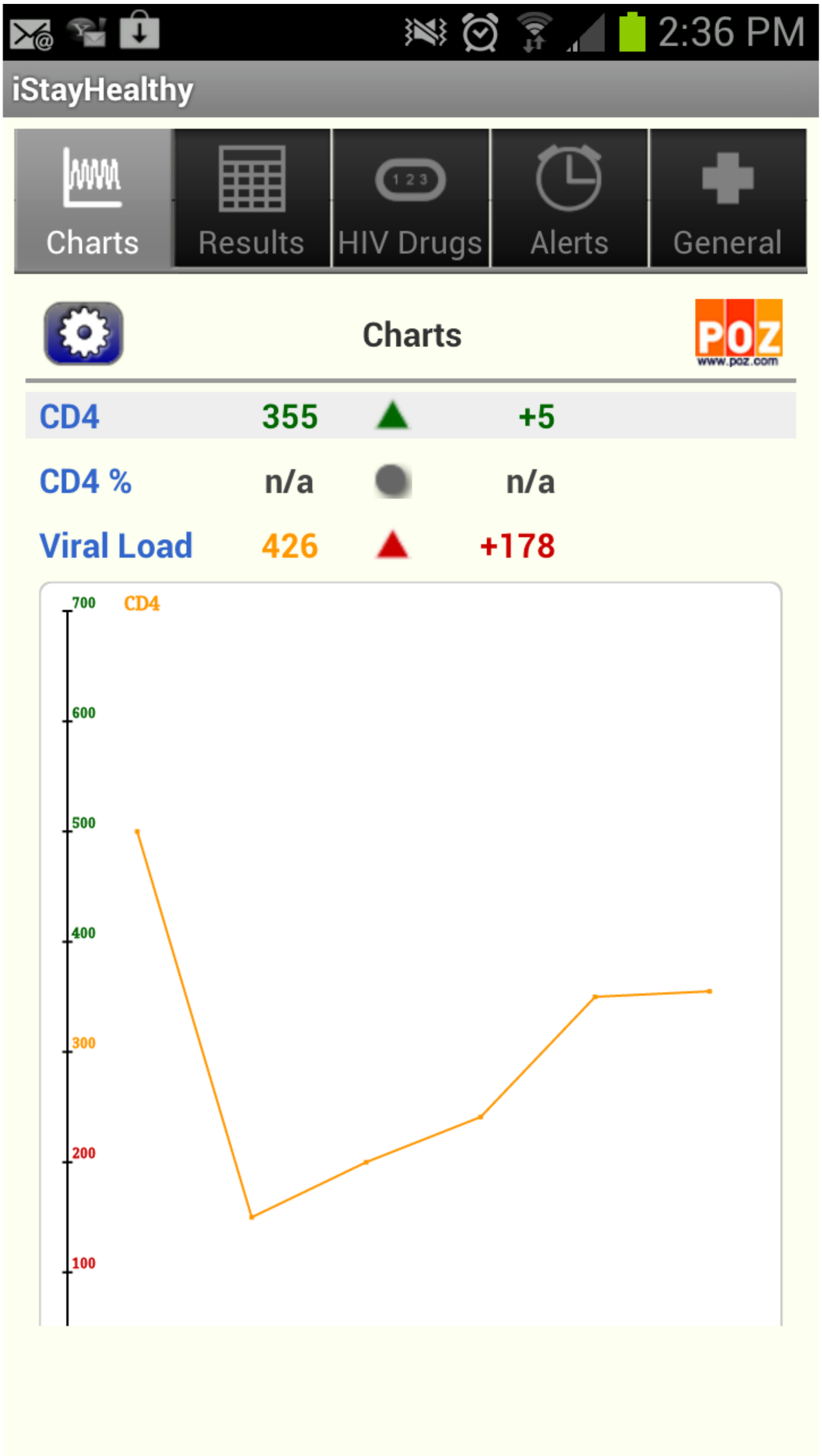
Out of the 55 apps, 8 (15%) were developed for persons diagnosed with HIV or other STDs. These apps typically included interactive tools for the following areas: medication adherence tracking and reminders, medication interaction information, medical appointment calendars and reminders, doctor/clinic names and locations, symptoms/side effects trackers, and viral load/lymphocyte (CD4) cell count trackers (Table 3). One social networking app (PositiveSingles) [29] facilitated connections with other persons diagnosed with HIV or other STDs.

Table 3. Apps for HIV/STD-positive persons.

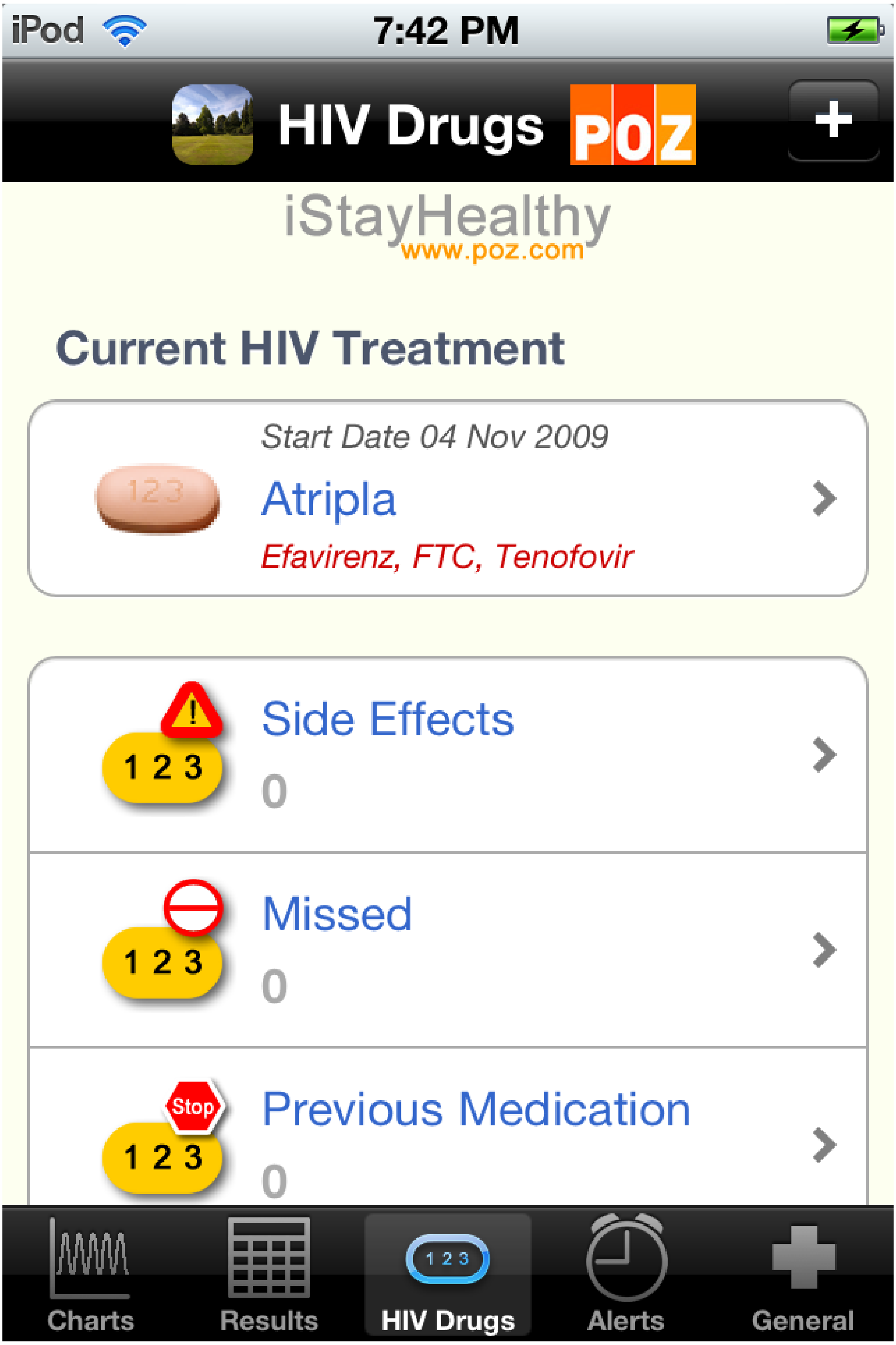
Name	Platform	Price (US $)	Features
Birmingham AIDS Outreach	iPhone, Android	Free	Connect to resources and support groups
HIV and Your Heart	iPhone	Free	Track CD4, viral load, weight, smoking, and other heart health measures
iStayHealthy	iPhone, Android	Free	Track CD4, viral load, medications, adherence, side effects, illnesses; alarm
Patient Treatment Companion	iPhone, Android	Free	Track CD4, viral load, medications, appointments, weight; alarm
PositiveSingles	iPhone, Android	Free	Connect to other positive persons
PozTracker	iPhone	3.99	Track CD4, viral load, medications, clinic locations, allergies, insurance; alarm
Red Ribbon HIV/AIDS Manager	iPhone, Android	9.99, 1.99	Record allergies, medications, CD4, viral load; alarm
Talk Positive	iPhone	Free	Track side effects, CD4, viral load, medications, appointments, allergies

The two highest user-rated apps for HIV/STD-positive persons were HIV and Your Heart [30] (5 stars) and iStayHealthy [31] for both the iPhone and Android (4.7 stars). The HIV and Your Heart [30] app included videos featuring providers, experts, and HIV-positive persons. This app also provided information about HIV and heart health, smoking cessation, diet and exercise, and a wellness checklist for tracking 9 health measures, including viral load and CD4 cell counts. The iStayHealthy [31] app provided a tracking and charting feature for viral load and CD4 cell counts (Figure 2), a component to record missed medication doses and medication side effects, a medication reminder alarm, and a place to record other medical history data, such as clinic locations and illness events. This app also included a photo library of HIV medication pill images that could be tailored to a daily medication schedule (Figure 3).

Figure 2. Screenshot of the iStayHealthy app’s user-generated CD4 cell count tracking. Android version shown on Samsung Galaxy S IV. Actual phone screen size: 9.6 × 5.5 cm. Reproduced with permission from creator, Dr Peter Schmidt.



CD4	355	▲	+5
CD4 %	n/a	●	n/a
Viral Load	426	▲	+178



Current HIV Treatment

Start Date 04 Nov 2009

Atripla

Efavirenz, FTC, Tenofovir

Side Effects

0

Missed

0

Previous Medication

0

Figure 3. Screenshot of iStayHealthy app's user-generated HIV medication list. iPhone version shown on Apple iPhone 4. Actual phone screen size: 6.8 × 4.5 cm. Reproduced with permission from creator, Dr Peter Schmidt.

Interactive App Components

Fifteen apps (27%) that did not specifically target HIV-infected persons included an interactive component, such as a game/quiz (n=9) or a risk assessment activity (n=6). Here we highlight 4 examples of different types of interactive components (full descriptions of all interactive app components available upon request). In Condom Pro [32], players advance through 10 game levels by using the phone's touchscreen to carefully "open" a condom, correctly place it over an object (Figure 4), and correctly remove it from the object. It's Your Shout [33] asks the user a series of humorous personality questions and then matches the user's profile with a superhero or celebrity's profile. The final message with each quiz announces that, "Even superheroes can get Chlamydia. Be tested. Be sure." The user is then given the option to order a home testing kit for chlamydia screening. The STD411 [24] app features an interactive chart for gay, bisexual, and transgender men where flashing colored condoms (red, yellow, and green) indicate the level of risk of different sexual activities for various STDs (Figure 5). As the user selects various colored condoms, diseases, or sexual activities, a brief risk explanation appears above the graphic. Finally, SeX Factor [34] includes a game in which the user moves across a game board by choosing and correctly answering an HIV/STD quiz question. Users who successfully cross the game board win a free chlamydia home test kit.

Figure 4. Screenshot of user placing a condom on a practice object in the interactive Condom Pro app. iPhone version shown on Apple iPad 2 for clarity. Reproduced with permission from Liz Sabatiuk, Social Media Manager, The National Campaign to Prevent Teen and Unplanned Pregnancy.

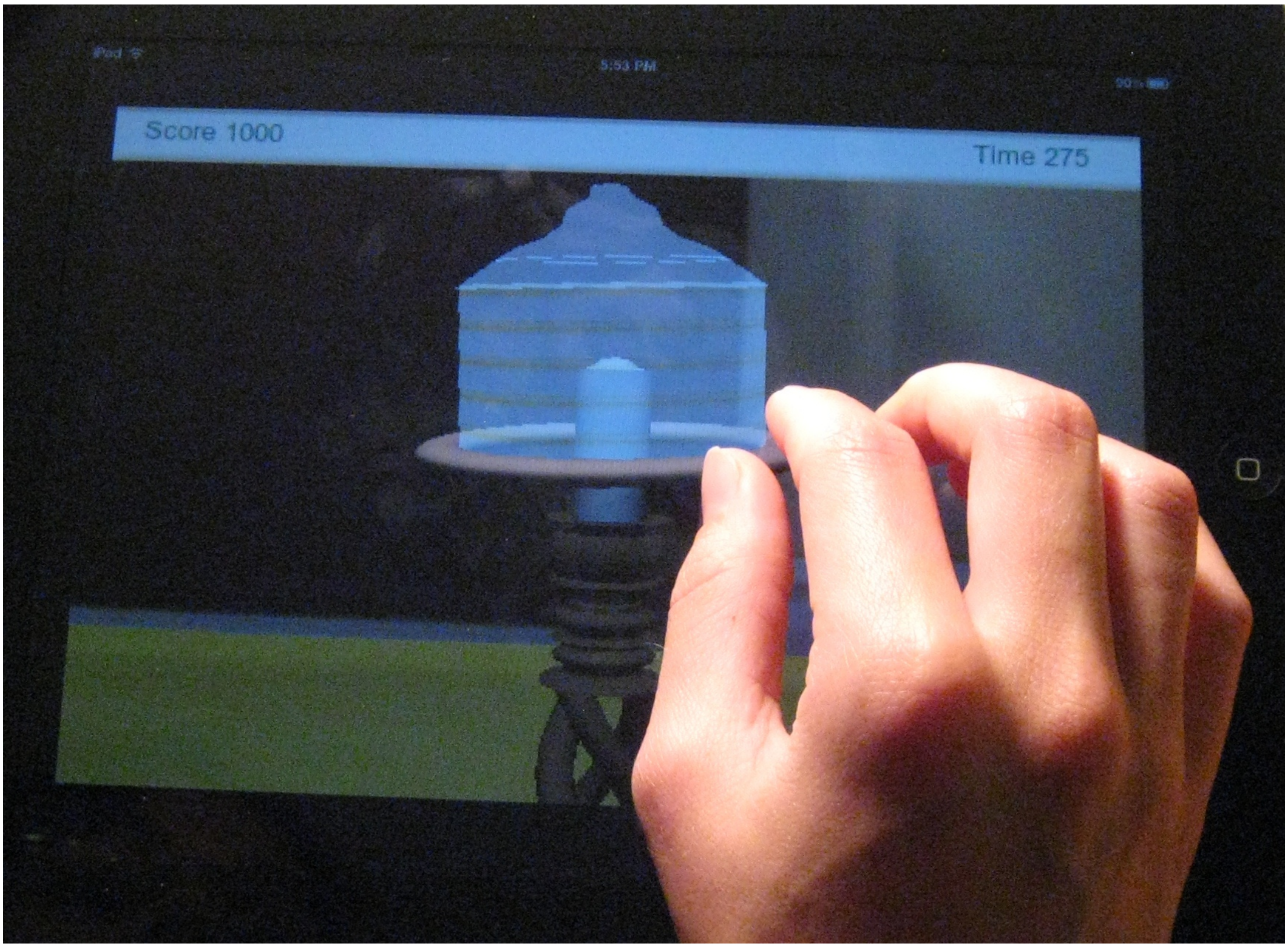


Figure 5. Screenshot of the MSM-tailored, interactive sexual behavior risk chart from STD411. iPhone version shown on Apple iPad 2 for clarity.

You can't give Syphilis back to yourself, but your buddy's finger or other body part can expose you if it has a sore on it.

App Ratings

Customer ratings and number of app downloads roughly measure an app's acceptability and popularity. More than half (38/69, 55%) of the apps were customer rated yielding an average store rating of 3.7 out of 5 stars. In general, apps were infrequently downloaded (median 100-500 downloads) with only 11 apps exceeding 1000 downloads. Most (46/69, 67%) of the apps were free; the remainder ranged in price from US $0.99 to $9.99. Among apps that charged a fee, those charging over the median of US $0.99 all received fewer than 100 downloads.

The most popular app based on reported number of downloads was Sex Facts [35] that had a star rating of 4.0 and downloads between 500,000 and 1,000,000. This app featured a revolving index of sex-related factoids that could be shared with friends via social network sites (eg, Facebook and Twitter) or email. The included HIV/STD facts on this app were not comprehensive and it was not clear how the specific facts were chosen. For example, only 2 facts referred to HIV transmission/acquisition: “HIV is transmitted through the blood” and “According to the Centers for Disease Control and Prevention (CDC), using a condom every time you have sex can lower your risk of HIV infection.” Regarding other sexually transmitted diseases, only human papillomavirus (HPV) was mentioned by name in a factoid that stated that HPV “is the world’s most common sexually transmitted infection.” Although sexuality was listed in the app’s description, the overwhelming majority of facts centered on heterosexuality.

Apps Tailored for Specific Populations

A small proportion of apps appeared to include or target specific populations. Approximately one-quarter (13/55, 24%) of apps were designed for a particular geographic area: England (Big Night Out [36], It’s Your Shout [33], Protection-Sex [37], Safer Sex [38], Sex Guide [39]); China/Singapore (Action For AIDS [22], My First Time [40]), Denmark (Safesex Guide [41]); Ireland (Sexual Health Guide [27]); Birmingham, Alabama (Birmingham AIDS Outreach [42]); Massachusetts (MASTDinfo [43]), New York City (NYC Condom Finder [44]); Philadelphia, Pennsylvania (iCondom Philly [45]); and San Francisco, California (STD411 [24]).

A minority of apps (n=10) were inclusive of information about anal sex or featured information for lesbian, gay, bisexual, or transgender (LGBT) persons. These apps included Birmingham AIDS Outreach [42], HIV RISK Calculator [46], MASTDinfo [43], Safe Sex Tips [26], Safer Sex [38], Sex Detective [47], Sex Guide [39], Sexual Health Guide [27], STD Risk Calculator Lite [48], and STD411 [24]. Of these, the only tailored apps were MASTDinfo [43], which provided information and testing center locations for men who have sex with men (MSM) in Massachusetts, and STD411 [24] that provided information and resources for gay, bisexual, and transgender men in San Francisco (Figure 5).

Nine apps included pictures, text, or videos with persons from racial or ethnic minority groups: AfterSex [49], AIDSinfo HIV/AIDS Glossary [50], Birmingham AIDS Outreach [42], HIV and Your Heart [30], Kenny Condom [51], Men’s Sexual Health [52], Pos or Not [23], STDiQ [25], and TKNO [53]. Among these, STDiQ [25] was tailored for black and Latino men and women. This app features videos from the Safe in the City campaign (Figures 6 and 7), a theory-based intervention found to reduce incident STD cases [54]. Kenny Condom [51] is based on a previously developed character and Web initiative tailored toward young African American men and “the hip-hop generation.”

Figure 6. Screenshot of STDiQ app featuring racially/ethnically diverse videos from the Safe in the City STD reduction intervention. iPhone version shown on iPad 2 for clarity. Reproduced with permission from Cornelis A Rietmeijer, MD, PhD, MSPH, University of Colorado, Denver.

Safe in the City is a 23-minute video that has been shown to reduce new STDs among STD clinic patients. Learn more at safeinthecity.org.

Figure 7. Screenshot of STDiQ app featuring a bar scene from Rueben's one-night stand in the Safe in the City STD reduction intervention. iPhone version shown on iPad 2 for clarity. Reproduced with permission from Cornelis A Rietmeijer, MD, PhD, MSPH, University of Colorado, Denver.

Discussion

In our review of Apple iTunes and Android Google Play apps, we found 55 unique apps that fostered HIV/STD prevention and care services. In general, these apps were underutilized and only a few apps received high user ratings. Apps primarily focused on providing disease information, and only 6 apps covered all 4 areas of a basic prevention strategy (knowledge raising, risk reduction, condom promotion, and testing). However, we also found examples of apps that used a tailored and/or interactive approach and apps that illustrated the potential for integrating mobile phone technologies into more comprehensive HIV prevention and care.

There are a number of benefits to using mobile phone apps to provide HIV/STD prevention and care services. Apps offer convenience to both the user and the developer because they provide a flexible way to reach a large audience at an affordable cost. Apps can provide individually tailored and interactive HIV/STD prevention interventions that are constantly accessible and allow the user to seek information while maintaining anonymity. Intervention customization and interactivity have been found to be important for effectiveness in behavior change interventions, including those that are technology based [55]. Well-designed mobile phone apps accommodate self-tailoring and personalization by allowing the user to choose which app features they wish to use and selecting messages and notifications they wish to receive through the app, email, or SMS. Unlike SMS and automated voice messages that are sent at specific times, the interactive functions of mobile phone apps are constantly available, allowing users to engage with the app at their convenience.

The HIV/STD phone apps also capitalize on delivering an intervention in a way that is familiar, desirable, and discrete to at-risk populations including youth, MSM, and racial and ethnic minorities [56]. Within the United States, the black population represents the most active and fastest growing user group of the mobile Internet [57]. And MSM have also been shown to have greater access to and use of cell phone technologies compared to heterosexual populations [58], increasing the likelihood of delivering highly engaging prevention messages to this hard-to-reach population, including those who are not reached through in-person or group interventions [59].

Unfortunately, many of the very features that make apps such a promising platform for delivering HIV/STD prevention and care services appear to be lacking in the currently available HIV/STD-related apps. Based on our review, as of August 1, 2012, less than 0.3% of the more than 29,000 health-related apps available for iPhone and Android consumers [18,19] were dedicated to HIV/STD information and prevention. Furthermore, existing HIV/STD apps have generally failed to attract the attention and positive reviews of target audiences. This is likely because of a combination of inadequate app promotion and failure to create engaging, attractive apps. How can we do better?

Many of the core elements of social marketing offer an appropriate guiding framework for developing health interventions with mobile technologies [60,61]. Limited downloads and low user ratings indicate that most HIV/STD apps have not achieved the ideal social marketing mix of “product, price, place, and promotion.” Social marketing principles, such as audience analysis and segmentation, and the use of formative research during product design would ensure that apps are developed with the user in mind [62]. For example, in focus groups we have held with young, black MSM in North Carolina, more men owned mobile phones that operated on an Android OS rather than an iOS platform (Muessig et al, unpublished). Thus, although a larger overall number of users in this age bracket may own iPhones, those who face a disproportionately high risk of HIV infection (young, black, and MSM) [63] might be better reached with an Android-supported app. Similarly, HIV/STD app developers need detailed information about who does and does not have access to these technologies. For example, general consumer research shows disparities in smartphone ownership by income. However, in our focus groups among young, black MSM, we found consistent smartphone ownership even among men with annual income under US $12,000 (Muessig et al, unpublished).

Relatedly, apps should include content and form that resonates with specific populations and that they want or need to use [64]. For example, an app that features a 3-paragraph narrative description about HIV clinic visits could be transformed into an interactive 10-point checklist of questions and topics that a patient could use to prepare for the appointment. This altered format engages the user and could be designed to allow users to self-tailor by entering their own checklist items.

Social marketing approaches also incorporate systems for process evaluation and iterative feedback. Ideally, app developers should be collecting target audience input throughout the development and usability testing process, as well as regularly updating their apps based on user feedback and advances in the HIV/STD field. Updates are especially critical for apps that provide information about rapidly evolving topics, such as testing locations, support services, and available HIV medications. As with any other intervention, these apps should be evaluated for measures of effectiveness. This requires identifying outcome goals and measurement metrics at the start of app design and development (eg, increased medication adherence, increased condom use, increased HIV testing frequency). Successful technology interventions such as these will require more involved collaborations between developers and public health HIV/STD researchers and practitioners [65].

Market research shows that smartphone owners regularly remove unused apps from their phones, thus design and content must be sufficiently useful and entertaining for consumers to keep them on their phones [66]. One analysis of app retention (as defined by the use of an app in the past 7 days) found a global retention rate of only 14.8% and US rates at a lower 12% [67]. Apps that incorporate more interactive components, such as games, quizzes, and activities, may be able to improve attractiveness and user retention. The categories of “entertainment” and “games” comprise the largest market share of available, active apps (27% to 29%) [19,68] and half of the top 10 most-downloaded apps of 2011 were gaming apps [69]. Game features of health interventions (eg, health-related challenges and rewards, ability to “level up,” and use of avatars) have been shown to promote and sustain healthy behaviors [70-73]. Thus, building in more interactive and gaming features in HIV/STD apps could improve their attractiveness and chances for successful intervention.

In addition to a need for better tailoring, design, and interaction, there were a number of content areas missing from most apps. First, apps for HIV-positive persons generally lacked psychological and emotional support resources and messages about the prevention of STDs and onward HIV transmission. Second, additional tools are needed to promote HIV medication adherence, support clinical monitoring, and facilitate patient-provider relationships. Active technology tools for self-monitoring have facilitated a variety of health improvement programs from increases in physical activity [74] to depression management [75]. Interventions using electronic drug monitoring devices suggest that these tools could also provide benefits for HIV-positive persons in improving medication adherence [76,77]. A few of the apps we identified included these types of tools, and assessments are needed to identify optimal design and efficacy. Additional mechanisms should be tested for providing linkage between the user/patient and their health care team. These features could include remote health coaching, symptom and side effect monitoring, and provision of real-time feedback. Third, a stronger emphasis is needed on interpersonal skills in the context of HIV/STD risk. Very few apps modeled condom negotiation, HIV/STD status disclosure, or sexual decision making in the context of relationships (see STDiQ [25] and Pos or Not [23] for notable exceptions). Finally, there were almost no descriptions of biomedical HIV prevention options, such as earlier initiation of antiretroviral therapy or preexposure and postexposure prophylaxis.

This review has several limitations. Our searches were restricted to the Apple iTunes and Google Play app stores. We selected these phone platforms (iPhone and Android) because they account for over 85% of the global app market [78,79] and, thus, are likely to be representative of the apps available to the majority of smartphone users. However, there may be HIV/STD-related apps available through other platforms that were not captured in our review. In addition, the field of app development is in constant motion with large numbers of new apps being created and other apps being deleted every day. Illustrating the speed of this development, the number of available apps in the 4 leading app stores (Apple iTunes Store, Google Play Store, Windows Phone Store, and Blackberry App World) increased from 989,476 in December 2011 [79] to 1,574,645 in September 2012 [78]. Importantly, our review was not designed to evaluate the effectiveness of identified apps in achieving positive prevention or treatment outcomes. Furthermore, we were not able to assess the user profiles of those who downloaded these apps. Although a growing body of evidence suggests high acceptability and smartphone ownership across diverse users [62,80], as discussed previously, detailed sociodemographic information (eg, user age, education, socioeconomic class, sexual orientation, race/ethnicity, and native language) should critically inform the tailoring and targeting of HIV/STD prevention and care apps.

Providing HIV/STD prevention and care services through mobile phone apps shows great potential for growth, both in improving the acceptability and adoption of existing apps, and creating new HIV/STD apps. Future HIV/STD app development could be informed by the principles of social marketing to build appropriately tailored, interactive apps. As biomedical advances in antiretroviral treatment are bringing the prevention of onward HIV transmission within our reach, we can use the powerful, widespread technologies offered through mobile phone apps to explore behavioral interventions for risk reduction and close gaps in HIV/STD testing, treatment adherence, and retention in care.

Acknowledgments

The University of North Carolina at Chapel Hill’s Libraries provided support for open access publication. Kathryn E Muessig is supported by an NIH institutional training grant (5T32AI007001-35). Lisa Hightow-Weidman and Emily Pike are supported by the NIH (R01 MH093275-01). The views expressed herein do not reflect the official stance of any funding agencies.

Conflicts of Interest

None declared.

References

1. Lim MS, Hocking JS, Hellard ME, Aitken CK. SMS STI: a review of the uses of mobile phone text messaging in sexual health. Int J STD AIDS 2008 May;19(5):287-290. [doi: 10.1258/ijsa.2007.007264] [Medline: 18482956]
2. Mukund Bahadur KC, Murray PJ. Cell phone short messaging service (SMS) for HIV/AIDS in South Africa: a literature review. Stud Health Technol Inform 2010;160(Pt 1):530-534. [Medline: 20841743]
3. Ybarra ML, Bull SS. Current trends in Internet- and cell phone-based HIV prevention and intervention programs. Curr HIV/AIDS Rep 2007 Dec;4(4):201-207. [Medline: 18366952]
4. Cornelius JB, Cato M, Lawrence JS, Boyer CB, Lightfoot M. Development and pretesting multimedia HIV-prevention text messages for mobile cell phone delivery. J Assoc Nurses AIDS Care 2011;22(5):407-413. [doi: 10.1016/j.jana.2010.11.007] [Medline: 21256053]
5. Lim EJ, Haar J, Morgan J. Can text messaging results reduce time to treatment of Chlamydia trachomatis? Sex Transm Infect 2008 Dec;84(7):563-564. [doi: 10.1136/sti.2008.032441] [Medline: 18723584]
6. Cornelius JB, Cato MG, Toth JL, Bard PM, Moore MW, White A. Following the trail of an HIV-prevention Web site enhanced for mobile cell phone text messaging delivery. J Assoc Nurses AIDS Care 2012;23(3):255-259. [doi: 10.1016/j.jana.2011.03.002] [Medline: 21550826]
7. Lester RT, Ritvo P, Mills EJ, Kariri A, Karanja S, Chung MH, et al. Effects of a mobile phone short message service on antiretroviral treatment adherence in Kenya (WelTel Kenya1): a randomised trial. Lancet 2010 Nov 27;376(9755):1838-1845. [doi: 10.1016/S0140-6736(10)61997-6] [Medline: 21071074]
8. Hardy H, Kumar V, Doros G, Farmer E, Drainoni ML, Rybin D, et al. Randomized controlled trial of a personalized cellular phone reminder system to enhance adherence to antiretroviral therapy. AIDS Patient Care STDS 2011 Mar;25(3):153-161 [FREE Full text] [doi: 10.1089/apc.2010.0006] [Medline: 21323532]
9. Horvath T, Azman H, Kennedy GE, Rutherford GW. Mobile phone text messaging for promoting adherence to antiretroviral therapy in patients with HIV infection. Cochrane Database Syst Rev 2012;3:CD009756. [doi: 10.1002/14651858.CD009756] [Medline: 22419345]
10. da Costa TM, Barbosa BJ, Gomes e Costa DA, Sigulem D, de Fátima Marin H, Filho AC, et al. Results of a randomized controlled trial to assess the effects of a mobile SMS-based intervention on treatment adherence in HIV/AIDS-infected Brazilian women and impressions and satisfaction with respect to incoming messages. Int J Med Inform 2012 Apr;81(4):257-269. [doi: 10.1016/j.ijmedinf.2011.10.002] [Medline: 22296762]
11. Pop-Eleches C, Thirumurthy H, Habyarimana JP, Zivin JG, Goldstein MP, de Walque D, et al. Mobile phone technologies improve adherence to antiretroviral treatment in a resource-limited setting: a randomized controlled trial of text message reminders. AIDS 2011 Mar 27;25(6):825-834. [doi: 10.1097/QAD.0b013e32834380c1] [Medline: 21252632]
12. Dowshen N, Kuhns LM, Johnson A, Holoyda BJ, Garofalo R. Improving adherence to antiretroviral therapy for youth living with HIV/AIDS: a pilot study using personalized, interactive, daily text message reminders. J Med Internet Res 2012;14(2):e51 [FREE Full text] [doi: 10.2196/jmir.2015] [Medline: 22481246]
13. Swendeman D, Rotheram-Borus MJ. Innovation in sexually transmitted disease and HIV prevention: internet and mobile phone delivery vehicles for global diffusion. Curr Opin Psychiatry 2010 Mar;23(2):139-144 [FREE Full text] [doi: 10.1097/YCO.0b013e328336656a] [Medline: 20087189]
14. Apple Inc. Apple Press Info. 2012. Apple's App Store Downloads Top 25 Billion URL: http://www.apple.com/pr/library/2012/03/05Apples-App-Store-Downloads-Top-25-Billion.html [accessed 2012-03-05] [WebCite Cache ID 69ZjQGhWP]

15. NielsenWire. Nielsen News. 2012. Two Thirds of New Mobile Buyers Now Opt for Smartphones URL: http://blog.nielsen.com/nielsenwire/online_mobile/two-thirds-of-new-mobile-buyers-now-opting-for-smartphones/ [accessed 2012-10-09] [WebCite Cache ID 6BI4ozpDo]
16. Joint United Nations Programme on HIV/AIDS (UNAIDS). UNAIDS. Geneva: United Nations; 2012 Jul 18. Together We Will End AIDS URL: http://www.unaids.org/en/media/unaids/contentassets/documents/epidemiology/2012/JC2296_UNAIDS_TogetherReport_2012_en.pdf [accessed 2012-08-07] [WebCite Cache ID 69kPezYEt]
17. Tech Computer Science. 2012 May 07. Google Play Reaches 15 Billion Downloads URL: http://techcomputerscience.com/blog/2012/05/07/google-play-reaches-15-billion-downloads/ [accessed 2012-07-31] [WebCite Cache ID 69ZjrnCgB]
18. Apple Inc. iTunes App Store. 2012. URL: https://www.sugarsync.com/pf/D614487_3721237_901149 [accessed 2012-08-01] [WebCite Cache ID 69b04X4oh]
19. AppBrain. 2012. Most popular Android market categories URL: http://www.appbrain.com/stats/android-market-app-categories [accessed 2012-07-31] [WebCite Cache ID 69ZkXOzu6]
20. The World Health Organization (WHO). Global health sector strategy on HIV/AIDS 2011-2015. Genva, Switzerland: The World Health Organization; 2011. URL: http://whqlibdoc.who.int/publications/2011/9789241501651_eng.pdf [accessed 2012-08-01] [WebCite Cache ID 69kE2t23a]
21. The White House Office of National AIDS Policy. National HIV/AIDS Strategy for the United States. Washington, DC; 2010 Jul 13. URL: http://www.whitehouse.gov/sites/default/files/uploads/NHAS.pdf [accessed 2012-08-07] [WebCite Cache ID 69kDUucak]
22. Codigo Pte Ltd. iTunes. Action for AIDS URL: https://itunes.apple.com/sg/app/action-for-aids/id417195119?mt=8 [accessed 2012-12-18] [WebCite Cache ID 6D0KRTUAg]
23. MTV Networks. iTunes. Pos or Not URL: https://itunes.apple.com/en/app/pos-or-not/id379498962?mt=8 [accessed 2012-12-18] [WebCite Cache ID 6D0KgmTLG]
24. Strona F. iTunes. STD411 URL: https://itunes.apple.com/us/app/std411/id395296278?mt=8 [accessed 2012-12-18] [WebCite Cache ID 6D0KxSlD7]
25. Denver Health. iTunes. STDiQ URL: https://itunes.apple.com/ca/app/stdiq/id390940551?mt=8 [accessed 2012-12-18] [WebCite Cache ID 6D0LSmNEH]
26. Solanki D. iTunes. Safe Sex Tips URL: https://itunes.apple.com/vg/app/safe-sex-tips/id366289353?mt=8 [accessed 2012-12-18] [WebCite Cache ID 6D0LgPLOj]
27. GIRT Mobile. Google Play. Sexual Health Guide URL: https://play.google.com/store/apps/details?id=com.grabradioworld.sexualhealthguide&hl=en [accessed 2012-12-18]
28. STFree Certifications Inc. iTunes. Chec-Mate URL: https://itunes.apple.com/us/app/chec-mate/id501324174?mt=8 [accessed 2012-12-18] [WebCite Cache ID 6D0MAr7Si]
29. SuccessfulMatch.com. Google Play. PositiveSingles URL: https://play.google.com/store/apps/details?id=com.successfulmatch.positivesingles&hl=en [accessed 2012-12-18]
30. Left Brain Media. iTunes. HIV and Your Heart URL: https://itunes.apple.com/us/app/hiv-and-your-heart/id423243195?mt=8 [accessed 2012-12-18] [WebCite Cache ID 6D0Mi9WTZ]
31. Schmidt P. iTunes. iStayHealthy URL: https://itunes.apple.com/us/app/istayhealthy/id422663813?mt=8 [accessed 2012-12-18] [WebCite Cache ID 6D0N0eBVD]
32. Bedsider LLC. iTunes. Condom Pro URL: https://itunes.apple.com/us/app/condom-pro/id436394146?mt=8 [accessed 2012-12-18] [WebCite Cache ID 6D0buCaGC]
33. NHS Blackpool. iTunes. Its Your Shout URL: https://itunes.apple.com/us/app/its-your-shout/id486778590?mt=8 [accessed 2012-12-18] [WebCite Cache ID 6D0cBt8P0]
34. ICE. iTunes. SeX Factor URL: https://itunes.apple.com/gb/app/sex-factor/id418132844?mt=8 [accessed 2012-12-18] [WebCite Cache ID 6D0cOtHJi]
35. RoidApp. Google Play. Sex Facts URL: https://play.google.com/store/apps/details?id=com.roidapp.sexfacts&hl=en [accessed 2012-12-18]
36. C-Dimension Ltd. iTunes. Big Night Out URL: https://itunes.apple.com/ca/app/ihit-big-night-out/id369547774?mt=8 [accessed 2012-12-18] [WebCite Cache ID 6D0czEaFW]
37. C-Dimension Ltd. Google Play. Protection-Sex URL: https://play.google.com/store/apps/details?id=com.cdimension.hit.protectionsex&hl=en [accessed 2012-12-18]
38. C-Dimension Ltd. Google Play. 2012 Dec 18. Safer Sex URL: https://play.google.com/store/apps/details?id=com.cdimension.hit.safersex&hl=en [accessed 2012-12-18]
39. Sexodroid. Google Play. Sex Guide URL: https://play.google.com/store/apps/details?id=com.sexodroid.sexguide [accessed 2012-12-18] [WebCite Cache ID 6D0ddcREJ]
40. Juzz Web Pte Ltd. iTunes. My First Time URL: https://itunes.apple.com/pk/app/my-first-time/id370386302?mt=8 [accessed 2012-12-18] [WebCite Cache ID 6D0eB1bbe]
41. Mobile Identity K/S. iTunes. Safesex Guide URL: https://itunes.apple.com/ca/app/safesex-guide/id411641559?mt=8 [accessed 2012-12-18] [WebCite Cache ID 6D0eMqMD2]

42. BAO. iTunes. Birmingham AIDS Outreach URL: https://itunes.apple.com/us/app/bao-for-iphone/id520093445?mt=8 [accessed 2012-12-18] [WebCite Cache ID 6D0easjtW]
43. Fenway Health. iTunes. MASTDinfo URL: https://itunes.apple.com/kz/app/mastdinfo/id407553492?mt=8 [accessed 2012-12-18] [WebCite Cache ID 6D0epbDr3]
44. Matsuki JM. iTunes. NYC Condom Finder URL: https://itunes.apple.com/us/app/nyc-condom-finder-by-nyc-health/id418902795?mt=8 [accessed 2012-12-21] [WebCite Cache ID 6D4pIWkaX]
45. Raaza. iTunes. iCondom Philly URL: https://itunes.apple.com/us/app/icondom-philly-philadelphias/id428303918?mt=8 [accessed 2012-12-21] [WebCite Cache ID 6D4oh09mo]
46. Aedes Labs. Google Play. HIV RISK Calculator URL: https://play.google.com/store/apps/details?id=aed.hivrisk [accessed 2012-12-21] [WebCite Cache ID 6D4pZ7GyE]
47. HRH Innovations. Google Play. Sex Detective URL: https://play.google.com/store/apps/details?id=com.sexdetectiveapp.sexdetectivefull&hl=en [accessed 2012-12-21]
48. Universal Web Solutions CA. Google Play. STD Risk Calculator Lite URL: https://play.google.com/store/apps/details?id=com.uwsca.STDRiskCalculatorLite&hl=en [accessed 2012-12-21]
49. Sachem Software LLC. Google Play. AfterSex URL: https://play.google.com/store/apps/details?id=com.sachem.aftersexhealth&hl=en [accessed 2012-12-21]
50. National Library of Medicine. iTunes. AIDSinfo HIV/AIDS Glossary URL: https://itunes.apple.com/ca/app/aidsinfo-hiv-aids-glossary/id397417517?mt=8 [accessed 2012-12-21] [WebCite Cache ID 6D4qOZ9pW]
51. Payton L. Google Play. Kenny Condom URL: https://play.google.com/store/apps/details?id=com.wKennyCondom&hl=en [accessed 2012-12-21]
52. Medicon Apps. iTunes. Men's Sexual Health URL: https://itunes.apple.com/us/app/mens-sexual-health/id326139389?mt=8 [accessed 2012-12-21] [WebCite Cache ID 6D4uCJLxo]
53. defineLabs. iTunes. TKNO URL: https://itunes.apple.com/gb/app/tkno/id412848909?mt=8 [accessed 2012-12-21] [WebCite Cache ID 6D4sO53mT]
54. Warner L, Klausner JD, Rietmeijer CA, Malotte CK, O'Donnell L, Margolis AD, Safe in the City Study Group. Effect of a brief video intervention on incident infection among patients attending sexually transmitted disease clinics. PLoS Med 2008 Jun 24;5(6):e135 [FREE Full text] [doi: 10.1371/journal.pmed.0050135] [Medline: 18578564]
55. Noar SM. Computer technology-based interventions in HIV prevention: state of the evidence and future directions for research. AIDS Care 2011 May;23(5):525-533. [doi: 10.1080/09540121.2010.516349] [Medline: 21287420]
56. Wise J, Operario D. Use of electronic reminder devices to improve adherence to antiretroviral therapy: a systematic review. AIDS Patient Care STDS 2008 Jun;22(6):495-504. [doi: 10.1089/apc.2007.0180] [Medline: 18462071]
57. Pew Internet. Two-thirds of young adults and those with higher income are smartphone owners. Washington, DC: Pew Internet & American Life Project; 2012 Sep 11. URL: http://pewinternet.org/~/media//Files/Reports/2012/PIP_Smartphones_Sept12%209%2010%2012.pdf [accessed 2012-10-11] [WebCite Cache ID 6BL7SWcmP]
58. LGBT Market Research and Development Lab. The LGBT Community Survey. San Francisco, CA: Community Marketing Inc; 2012 Jul. URL: http://www.communitymarketinginc.com/documents/CMI_6th_LGBT_Community_Survey_USv1.pdf [accessed 2012-10-10] [WebCite Cache ID 6BJhmVzwt]
59. Elford J, Bolding G, Davis M, Sherr L, Hart G. Web-based behavioral surveillance among men who have sex with men: a comparison of online and offline samples in London, UK. J Acquir Immune Defic Syndr 2004 Apr 1;35(4):421-426. [Medline: 15097159]
60. Lefebvre C. Integrating cell phones and mobile technologies into public health practice: a social marketing perspective. Health Promot Pract 2009 Oct;10(4):490-494. [doi: 10.1177/1524839909342849] [Medline: 19809002]
61. Grier S, Bryant CA. Social marketing in public health. Annu Rev Public Health 2005;26:319-339. [doi: 10.1146/annurev.publhealth.26.021304.144610] [Medline: 15760292]
62. Ramanathan N, Swendeman D, Comulada WS, Estrin D, Rotheram-Borus MJ. Identifying preferences for mobile health applications for self-monitoring and self-management: Focus group findings from HIV-positive persons and young mothers. Int J Med Inform 2012 Jun 13. [doi: 10.1016/j.ijmedinf.2012.05.009] [Medline: 22704234]
63. Centers for Disease Control and Prevention. 2012 Jul. HIV in the United States: at a glance URL: http://www.cdc.gov/hiv/resources/factsheets/us.htm [accessed 2012-10-11] [WebCite Cache ID 6BL7ZHdGx]
64. Rabin C, Bock B. Desired features of smartphone applications promoting physical activity. Telemed J E Health 2011 Dec;17(10):801-803. [doi: 10.1089/tmj.2011.0055] [Medline: 22010977]
65. Kennedy CM, Powell J, Payne TH, Ainsworth J, Boyd A, Buchan I. Active assistance technology for health-related behavior change: an interdisciplinary review. J Med Internet Res 2012;14(3):e80 [FREE Full text] [doi: 10.2196/jmir.1893] [Medline: 22698679]
66. Firminger R. iOS & Android App User Retention, Months Since Acquisition (%).: Flurry Analytics; 2011 Feb. Using analytics to take your mobile games to the next level URL: http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=5&ved=0CFUQFjAE&url=http%3A%2F%2Fcasualconnect.org%2Fcontent%2Feurope%2F2012%2FRichard_Firminger_CCE_2012.

pptx&ei=FUgZUJS_DpKO8wT-woCwDg&usg=AFQjCNGiTvNb8IXlSJunCzZ_xs84i3HBCA [accessed 2012-03-01] [WebCite Cache ID 69axLrZSD]
67. Farago P. Flurry Blog.: Flurry Analytics; 2009 Oct 15. Global app retention: a tale of two economies URL: http://blog.flurry.com/bid/27138/Global-App-Retention-A-Tale-of-Two-Economies [accessed 2012-08-01] [WebCite Cache ID 69axtK59W]
68. 148Apps. App Store Metrics. 2012. Current active application count by category URL: http://148apps.biz/app-store-metrics/?mpage=catcount [accessed 2012-08-01] [WebCite Cache ID 69b1ijq9j]
69. Perez S. Tech Crunch. 2011 Dec 20. Distimo's Year-End Report Shows Why Developers Love iOS: iPhone 4x Android Revenue, iPad 2x URL: http://techcrunch.com/2011/12/20/distimos-year-end-report-shows-why-developers-love-ios-iphone-4x-android-revenue-ipad-2x/ [accessed 2012-08-01] [WebCite Cache ID 69b24dWzG]
70. Brown SJ, Lieberman DA, Germeny BA, Fan YC, Wilson DM, Pasta DJ. Educational video game for juvenile diabetes: results of a controlled trial. Med Inform (Lond) 1997;22(1):77-89. [Medline: 9183781]
71. Baranowski T, Baranowski J, Cullen KW, Marsh T, Islam N, Zakeri I, et al. Squire's Quest! Dietary outcome evaluation of a multimedia game. Am J Prev Med 2003 Jan;24(1):52-61. [Medline: 12554024]
72. Zurovac D, Sudoi RK, Akhwale WS, Ndiritu M, Hamer DH, Rowe AK, et al. The effect of mobile phone text-message reminders on Kenyan health workers' adherence to malaria treatment guidelines: a cluster randomised trial. Lancet 2011 Aug 27;378(9793):795-803 [FREE Full text] [doi: 10.1016/S0140-6736(11)60783-6] [Medline: 21820166]
73. Lieberman D. Interactive video games for health promotion: Effects on knowledge, self-efficacy, social support,health. In: Health Promotion and Interactive Technology: Theoretical Applications and Future Directions. Mahwah, NJ: Lawrence Erlbaum Associates; 1997:103-120.
74. Consolvo S, Klasnja P, McDonald DW, Avrahami D, Froehlich J, LeGrand L, et al. Flowers or a robot army? Encouraging awareness & activity with personal, mobile displays. Proceedings of the 10th International Conference on Ubiquitous Computing Sept 21-24; Seoul, South Korea. New York: ACM; 2008 Sep 21. URL: http://dub.washington.edu/djangosite/media/papers/UbiComp188-consolvo.pdf [accessed 2012-12-21] [WebCite Cache ID 6D5HSoYyC]
75. Kauer SD, Reid SC, Crooke AH, Khor A, Hearps SJ, Jorm AF, et al. Self-monitoring using mobile phones in the early stages of adolescent depression: randomized controlled trial. J Med Internet Res 2012;14(3):e67 [FREE Full text] [doi: 10.2196/jmir.1858] [Medline: 22732135]
76. Sabin LL, DeSilva MB, Hamer DH, Xu K, Zhang J, Li T, et al. Using electronic drug monitor feedback to improve adherence to antiretroviral therapy among HIV-positive patients in China. AIDS Behav 2010 Jun;14(3):580-589 [FREE Full text] [doi: 10.1007/s10461-009-9615-1] [Medline: 19771504]
77. Krummenacher I, Cavassini M, Bugnon O, Schneider MP. An interdisciplinary HIV-adherence program combining motivational interviewing and electronic antiretroviral drug monitoring. AIDS Care 2011 May;23(5):550-561. [doi: 10.1080/09540121.2010.525613] [Medline: 21271406]
78. Statista. 2012. Number of apps available in leading app stores as of September 2012 URL: http://www.statista.com/statistics/208599/number-of-apps-available-in-leading-app-stores/ [accessed 2012-10-19] [WebCite Cache ID 6BXYwh0SN]
79. PR Newswire. 2011 Dec 02. Mobile Apps Approaching Major Milestone of 1 Million Apps in Marketplace: Mobilewalla URL: http://www.prnewswire.com/news-releases/mobile-apps-approaching-major-milestone-of-1-million-apps-in-marketplace-mobilewalla-134891433.html [accessed 2012-10-19] [WebCite Cache ID 6BXYm1RE2]
80. Zickuhr K, Smith A. Digital Differences. Washington, DC: Pew Internet & American Life Project; 2012 Apr 13. URL: http://pewinternet.org/~/media//Files/Reports/2012/PIP_Digital_differences_041312.pdf [accessed 2012-10-11] [WebCite Cache ID 6BL7HMIo4]

Abbreviations

HIV: human immunodeficiency virus
HPV: human papillomavirus
MSM: men who have sex with men
OS: operating system
SMS: Short Message Service
STD: sexually transmitted disease

Edited by G Eysenbach; submitted 15.08.12; peer-reviewed by E Gabarron; comments to author 06.10.12; revised version received 15.10.12; accepted 16.10.12; published 04.01.13.

Please cite as:
Muessig KE, Pike EC, LeGrand S, Hightow-Weidman LB
Mobile Phone Applications for the Care and Prevention of HIV and Other Sexually Transmitted Diseases: A Review
J Med Internet Res 2013;15(1):e1
URL: http://www.jmir.org/2013/1/e1/
doi:10.2196/jmir.2301
PMID:23291245

Original Paper

A Smartphone-Based Intervention With Diaries and Therapist-Feedback to Reduce Catastrophizing and Increase Functioning in Women With Chronic Widespread Pain: Randomized Controlled Trial

Ólöf Birna Kristjánsdóttir[1,2], RN, MSc; Egil A Fors[3,4], MD, PhD; Erlend Eide[1,5]; Arnstein Finset[2], MApsych, PhD; Tonje Lauritzen Stensrud[2], MApsych; Sandra van Dulmen[6,7,8], MApsych, PhD; Sigrid Hørven Wigers[9], MD, PhD; Hilde Eide[1,8], RN, MApsych, PhD

[1]Institute of Nursing, Faculty of Health, Oslo and Akershus University College of Applied Sciences, Oslo, Norway

[2]Department of Behavioral Sciences in Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway

[3]Department of Psychiatry; National Competence Centre for Complex Symptom Disorders, St Olav University Hospital, Trondheim, Norway

[4]Department of Public Health and General Practice, Faculty of Medicine, General Practice Research Unit, Norwegian University of Science and Technology, Trondheim, Norway

[5]edesign, Tønsberg, Norway

[6]NIVEL (Netherlands Institute for Health Services Research), Utrecht, Netherlands

[7]Department of Primary and Community Care, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands

[8]Faculty of Health Sciences, Buskerud University College, Drammen, Norway

[9]Jeløy Kurbad Rehabilitation Centre, Moss, Norway

Corresponding Author:
Ólöf Birna Kristjánsdóttir, RN, MSc
Department of Behavioral Sciences in Medicine
Faculty of Medicine
University of Oslo
Post Box 1111 Blindern
Oslo, 0317
Norway
Phone: 47 22851020
Fax: 47 22851300
Email: kristjansdottir@gmail.com

Abstract

Background: Internet-based interventions using cognitive behavioral approaches can be effective in promoting self-management of chronic pain conditions. Web-based programs delivered via smartphones are increasingly used to support the self-management of various health disorders, but research on smartphone interventions for persons with chronic pain is limited.

Objective: The aim of this trial was to study the efficacy of a 4-week smartphone-delivered intervention with written diaries and therapist feedback following an inpatient chronic pain rehabilitation program.

Methods: A total of 140 women with chronic widespread pain who participated in a 4-week inpatient rehabilitation program were randomized into 2 groups: with or without a smartphone intervention after the rehabilitation. The smartphone intervention consisted of 1 face-to-face session and 4 weeks of written communication via a smartphone. Participants received 3 smartphone diary entries daily to support their awareness of and reflection on pain-related thoughts, feelings, and activities. The registered diaries were immediately available to a therapist who submitted personalized written feedback daily based on cognitive behavioral principles. Both groups were given access to a noninteractive website after discharge to promote constructive self-management. Outcomes were measured with self-reported questionnaires. The primary outcome measure of catastrophizing was determined using the pain catastrophizing scale (score range 0-52). Secondary outcomes included acceptance of pain, emotional distress, functioning, and symptom levels.

Results: Of the 140 participants, 112 completed the study: 48 in the intervention group and 64 in the control group. Immediately after the intervention period, the intervention group reported less catastrophizing (mean 9.20, SD 5.85) than the control group (mean 15.71, SD 9.11, *P*<.001), yielding a large effect size (Cohen's *d*=0.87) for study completers. At 5-month follow-up, the between-group effect sizes remained moderate for catastrophizing (Cohen's *d*=0.74, *P*=.003), acceptance of pain (Cohen's *d*=0.54, *P*=.02), and functioning and symptom levels (Cohen's *d*=0.75, *P*=.001).

Conclusions: The results suggest that a smartphone-delivered intervention with diaries and personalized feedback can reduce catastrophizing and prevent increases in functional impairment and symptom levels in women with chronic widespread pain following inpatient rehabilitation.

Trial Registration: Clinicaltrials.gov NCT01236209; http://www.clinicaltrials.gov/ct2/show/NCT01236209 (Archived by WebCite at http://www.webcitation.org/6DUejLpPY)

(J Med Internet Res 2013;15(1):e5) doi:10.2196/jmir.2249

KEYWORDS

Widespread Chronic Pain; Fibromyalgia; Self-management; Mobile phones; Internet; Cognitive Therapy; Catastrophization, Recurrence

Introduction

Chronic widespread pain is a common cause of suffering. An estimated 4% to 10% of the adult population experiences chronic widespread pain, ie, musculoskeletal pain lasting for more than 3 months not caused by an identifiable physical pathology [1-5]. This pain is often accompanied by other symptoms, including fatigue, sleep disturbance, and emotional distress [2]. A subgroup meets the criteria for fibromyalgia syndrome, where in addition to the chronic pain, the pain thresholds are reduced and tenderness in more than 10 of 18 specified trigger points is identified [2,3]. The development and maintenance of chronic widespread pain and fibromyalgia involve a complex dynamic process with biological, cognitive, and psychosocial factors. The cause or underlying mechanisms are still not clearly identified and no single cure is available. Maladaptive thoughts and feelings seem to play an important part in the negative spiral resulting in the maintenance of chronic pain [6]. Multidimensional rehabilitation, including physical exercise and cognitive behavioral therapy (CBT), is recommended as treatment [7-8]. A key element is self-management, eg, balancing activity and rest, stress management, emotion regulation, and doing appropriate physical exercises [6-10]. However, relapse of symptoms is not uncommon [8,11,12] because self-management can be challenging due to the nature of the symptoms. Few studies have examined home-delivered interventions that aim to support self-management of chronic pain following rehabilitation [11-13].

Pain Conditions and Web-Based Interventions

Internet-based interventions using cognitive behavioral approaches can be effective in promoting self-management of chronic pain conditions [14-16]. Web-based programs delivered through smartphones are increasingly used to support the self-management of various health disorders; however, research on smartphone interventions for patients with chronic pain is limited [17]. Among the advantages of using smartphones rather than the traditional personal computer s are their small size and mobility, making self-management support available to the user in most situations [17]. Diaries with questions intended to support awareness and reflection are made available on the phone and the registered information can be submitted to a website and made instantly available to a therapist. Feedback can be automatically delivered and tailored to the registered information to some extent, or it can be even more personalized by a therapist [18-20]. In a recent study, a panel of health care professionals and people experiencing chronic pain discussed characteristics of a successful Internet self-management program. Important features included assisting patients to be more aware of their patterns of behavior and psychological experience, supporting the pursuit of personal goals and values-based behavior, and by using a small and mobile device for real-time monitoring and response [21]. The number of pain self-management applications for smartphones has increased exponentially since 2009 [22]. In 2010, more than 90 applications offering support in the self-management of chronic pain were available in application stores. There is a need for research in this field because many applications seem to have been developed without the involvement of a health care professional and, to our knowledge, none have been tested in randomized trials [22].

Theoretical Model

Cognitive and emotional factors influence the pain experience [23]. Among the psychological constructs that can play an important role in the development and maintenance of chronic pain is catastrophizing [6,23,24]. Pain-related catastrophizing includes the tendency to ruminate about and magnify symptoms, to expect the worst, and to feel helpless regarding self-management [25]. Catastrophizing tends to discourage patients from committing to their valued behavior and it has consistently been found to predict distress and disability [6,26,27]. In rehabilitation, catastrophizing is targeted in a number of ways, such as with CBT and exercise programs [6,28]. However, interventions delivered in the patient's private environment, supporting awareness of maladaptive thoughts and feelings, and providing personalized feedback may further help reduce catastrophizing [13,18]. A mobile phone-delivered intervention with diaries and daily CBT-based feedback has been found to reduce catastrophizing thoughts in patients with irritable bowel syndrome and these effects were maintained at a 3-month follow-up [18].

Acceptance and Commitment Therapy (ACT) is a third-generation CBT based on the notion that suffering may largely be caused by thinking about painful experiences rather than the experiences themselves [29]. Suffering can be reduced through mindfulness, acceptance, and committed action [29]. ACT has been found to be effective for people with various chronic health disorders [30], and has been used successfully to reduce catastrophizing and disability in chronic pain patients [31-33]. The goals are to promote psychological flexibility, such as acceptance of, rather than struggling with, unwanted thoughts, emotions, and symptoms (eg, pain or catastrophizing) and to increase commitment to personal values [28,29,34]. A person's values are described as his or her desired way of being within various life domains (eg, being a caring friend). Values differ from goals in that they can never be fully obtained, but can give a continuous sense of motivation, direction, and purpose [28]. The focus on values is also evident in the self-determination theory (SDT) that states the importance of perceiving behavior as self-determined for intrinsic motivation to be maintained [35]. According to the SDT, context-specific feedback can play a role in enhancing intrinsic motivation to maintain behavior [35]. Guidance was also found in the elaboration likelihood model of persuasion theory [36]. This theory specifies how information can be constructed and presented to enhance either cognitive elaboration or emotional elaboration intending to influence behavior change. Elements focused on in this study are repetition, personal relevance and involvement influencing the cognitive level, and influencing emotional pathways through emotion recognition, mindfulness exercises, and empathic communication.

Aims of the Study

We hypothesized that receiving personalized feedback shortly after having registered pain-related thoughts, feelings, and self-management activities in an everyday setting might reduce catastrophizing and increase functioning. The results of our pretrial study of a similar smartphone intervention indicated feasibility and user-friendliness for patients with chronic widespread pain [20].

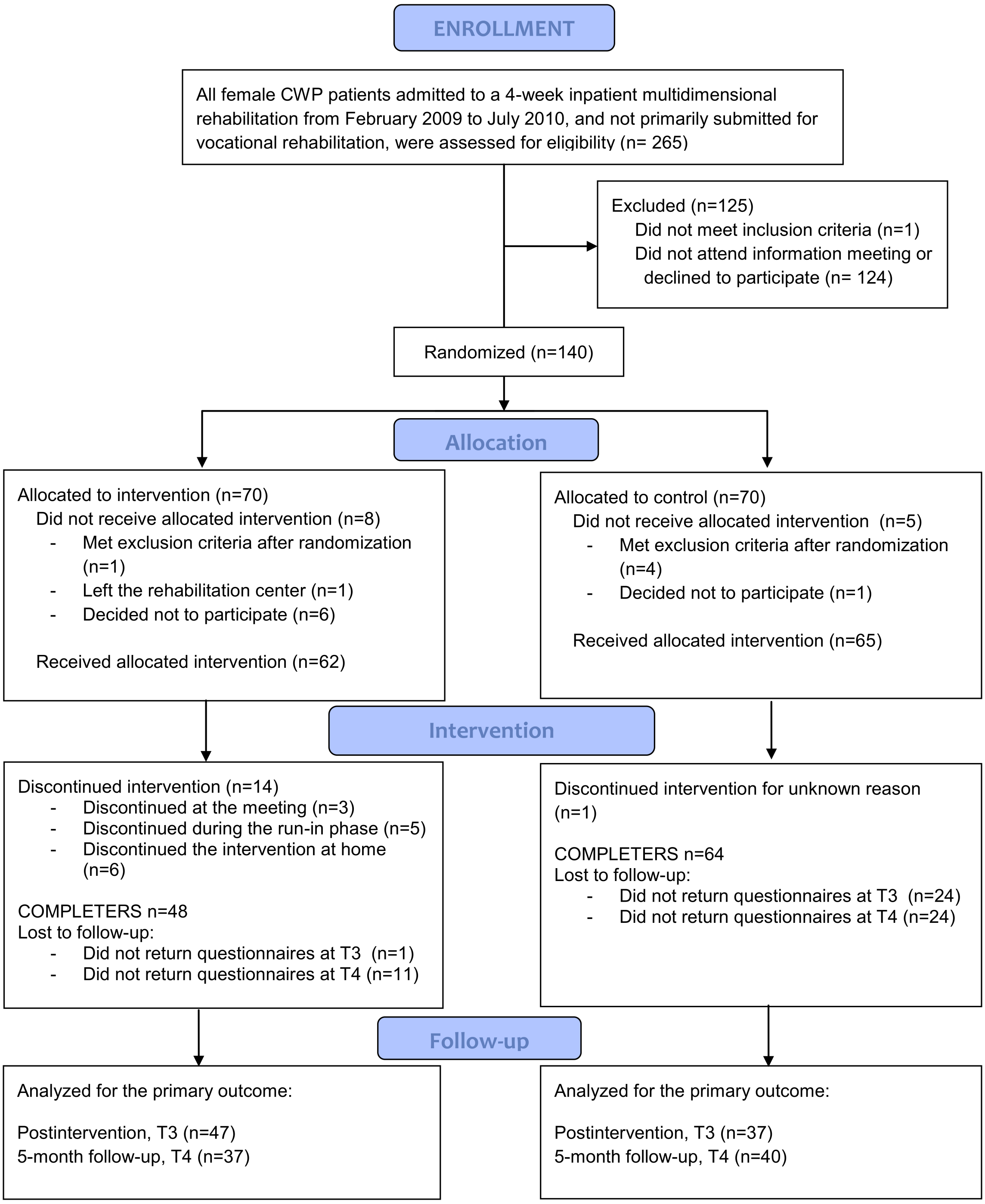
The present randomized controlled trial investigates the efficacy of a smartphone intervention on catastrophizing, acceptance, emotional distress, values-based behavior, and functioning and symptom level in women with chronic widespread pain who had completed a 4-week inpatient rehabilitation program. For the first 4 weeks after discharge, the intervention group received a Web-based intervention comprising registration of symptoms, thoughts, feelings, and self-management behavior through daily smartphone diaries and written personalized CBT-based feedback. It was hypothesized that the intervention group would show less catastrophizing and emotional distress, more acceptance of pain, and success in values-based living, and improved functioning and symptom levels after completing the intervention period and at a 5-month follow-up compared to a control group.

Methods

Study Design

The overall study design is shown in Figure 1. The design is a parallel-group, randomized controlled trial. Block randomization was used for practical reasons to ensure similar numbers in each group at each time point. All participants attended a 4-week inpatient multidimensional rehabilitation program for chronic pain (see Treatment Procedures). In the fourth week of the program, participants were randomly assigned to 1 of the 2 study groups. The intervention group received a smartphone intervention for 4 weeks after completing the inpatient rehabilitation. Both groups were given access to a noninteractive website with self-help pain management material. Self-reported assessments were gathered at 4 time-points: before (T1) and after (T2) the inpatient program, 4 weeks after discharge when the intervention group had completed their smartphone intervention (T3), and 6 months after discharge from the rehabilitation center (T4). The first 2 questionnaires were received and completed at the rehabilitation center and the last 2 were completed at home and returned by mail. One reminder letter was sent followed by a phone call from a researcher if the questionnaire was not returned.

Figure 1. Study design and participant flow.



Participants

Participants were recruited consecutively from Jeløy Kurbad Rehabilitation Center in Moss, Norway. Patients were referred to the center by their general practitioner or a medical specialist.

The inclusion criteria were: female, 18 years or older, participating in the inpatient multidimensional rehabilitation program for chronic pain, having chronic widespread pain for more than 6 months (with or without a diagnosis of fibromyalgia), not participating in another research project at the rehabilitation center, being able to use a smartphone, and not being diagnosed with a profound psychiatric disorder. The study took place between February 2009 and August 2010.

Ethical Aspects

The study was approved by the Regional Ethics Committee in South-East Norway and by the Norwegian Social Science Services. All participants signed an informed consent form. The study is registered at ClinicalTrial.gov (NCT01236209).

Procedures

At admission to the inpatient rehabilitation program, all chronic widespread pain patients received a written invitation to attend an informational group meeting where a researcher or a research assistant presented the study. Those who were interested in participating and met the inclusion criteria were given an informed consent form to sign.

A computer-generated sequence list with the 2 groups randomized in blocks of 4 was used because admission of 4 patients per week was expected. The computer-generated inclusion pattern was either 2 participants in each group or 3 to one group, sometimes 3 in the control group and other times 3 in the intervention group, until the final number of 140 was reached. A research assistant put the allocation information in sequentially numbered envelopes and sealed them. A researcher subsequently gave each participant a number and opened the matched envelope to reveal the group allocation. The information about group allocation was revealed to the participant at the inclusion meeting with a nurse in the final week of the inpatient program.

Assessment Measures

Participants completed self-administered questionnaires in paper format on arrival at the rehabilitation center (T1), at discharge (T2), immediately after the smartphone intervention (T3), and 6 months after discharge from the rehabilitation center (T4), which was 5 months after the smartphone intervention.

The pain catastrophizing scale (PCS) [25] was used to measure the primary outcome variable of the study, catastrophizing. It is a 13-item questionnaire with questions on helplessness, magnification, and rumination. Patients rate items on a scale from 0 (not at all) to 4 (all the time). The total score range for the PCS is 0 to 52, with higher scores reflecting higher degrees of catastrophizing. In our sample, the internal consistency was high on all assessments (Cronbach alpha range .892 to .942). As in prior research, scores greater than 24 were considered high [25,37].

The chronic pain acceptance questionnaire (CPAQ) [38] was used to measure acceptance. It is a 20-item self-report instrument developed to capture the extent of participation in daily activities despite pain and willingness to experience pain without trying to control, alter, or avoid it. It is scored on a 7-point Likert scale (0 = never true; 6 = always true) to give the total score (0-120). Higher scores reflect higher acceptance of pain and higher activities engagement. The reliability of the CPAQ has been established [38]. In our study, the Cronbach alpha coefficients were .814 to .910.

The questions from the 12-item General Health Questionnaire (GHQ) were used [39] with modified response alternatives. Responses to all items were given on the same 4-point scale (much less than usual, same as usual, more than usual, and much more than usual), but not on 2 scales as in the original. The questions measure changes in emotional distress over the previous couple of weeks. A bimodal scoring method was used (1 = symptom present more than usual; 0 = symptom present less than or as usual). Total score range is 0 to 12; indicating the number of symptoms present more than usual during the past 2 weeks. In the current study, the Cronbach alpha coefficients were .703 to .871.

The Chronic Pain Values Inventory (CPVI) is a 12-item measure of importance and success in living according to one’s own values in 6 domains (family, intimate relationships, friendship, work, health, and personal growth) [40]. Each item is rated on a scale from 0 to 5, with higher numbers indicating more importance or success. The mean success rating was used as a measure of values-based action (score range 0-5), as suggested by the authors [40]. In the present study, the Cronbach alpha coefficients for the success scale were .754 to .882.

The current levels (past couple of days) of pain, fatigue, and sleep disturbance were assessed on visual analog scales (VAS) from 0 (no pain/fatigue/sleep disturbance) to 100 (worst imaginable pain/fatigue/sleep disturbance) because these are cardinal symptoms of chronic widespread pain and fibromyalgia.

The original version of the Fibromyalgia Impact Questionnaire (FIQ) was used to measure the impact of fibromyalgia on functioning and symptom levels the past week. It consists of 10 questions with different response alternatives. One question includes 10 subitems related to the ability to perform activities of daily living. The response alternatives are given on a 4-point scale. The other questions enquire about general well-being, ability to work, and level of pain, fatigue, stiffness, and symptoms of anxiety and depression. Questions on symptom level are answered using a VAS from 0 to 100 (high symptom level). The score range is 0 to 100; higher scores indicate greater impairment [41]. The Cronbach alpha coefficients were .807 to .860.

The Short-Form Health Survey (SF-8) was also used to measure functioning. The SF-8 includes 8 items, scored on 5- or 6-point Likert scales, regarding level of functioning the past week. Summary measure scales for the mental health component and the physical component were obtained by using SF-8 Scoring Software 4.5 [42]. Scoring is standardized using the means and standard deviations from a survey from the general adult population in the United States (standardized mean 50, SD 10). Higher scores indicate better functioning; scores above 50 indicate functioning above the average in the US population. In the Norwegian version used in the present study, wording of response options for 2 items differed slightly from the original. In the original, the response alternatives for the item on role physical are none at all, a little bit, some, quite a lot, and could not do daily work. In our version, instead of “a little bit” the response was “very little.” In the original, the response alternatives for the mental health item are not at all, slightly, moderately, quite a lot, and extremely. In our version, “very little" was used instead of "slightly." The Cronbach alpha coefficients were .785 to .865 in the present study. Use of the noninteractive website was assessed with a self-report 4 weeks after discharge (T3) on how often the participant had visited the website.

Feasibility of the smartphone intervention was assessed with single questions postintervention (T3). For example, "I feel it has been a burden to participate in this intervention (to fill out diaries and receive feedback)" with a 5-point Likert scale (1 = agree completely; 5 = disagree completely).

Treatment Procedures

Inpatient Multidimensional Rehabilitation

All participants participated in a 4-week inpatient multidimensional rehabilitation program for patients with chronic pain. It included education in pain mechanisms and CBT-based pain management (approximately 20 hours), group sessions based on motivational interviewing (4 hours), various forms of aerobic exercise (outdoors, in the pool, and in the gym), stretching, and relaxation. In addition, individual myofascial pain treatment was given in accordance with the protocol of Travell [43,44] and medication was administered as needed (see [9] for details of the program).

Smartphone Intervention: Diaries and Daily Situational Feedback

The intervention was developed in 2008. One of the authors (EE) was responsible for the software development. The usability of the intervention was tested in a pretrial study with 6 women with chronic pain. Participation was experienced by the majority as supportive and motivating [20]. The key ACT concepts and a summary of their operationalization in the intervention are shown in Table 1 [28,34].

Table 1. Examples of Acceptance and Commitment Therapy (ACT) elements in diaries and feedback.

ACT element	Aim of diaries	Examples of diary questions	Aim of feedback	Examples of feedback
Cognitive defusion/ mindfulness	Awareness supported by making diary entries on thoughts, feelings and behavior three times a day	(1) Right now, my breathing is deep and relaxed. (2) Right now, I believe it is harmful for me to use my body. (3) Right now, I am coping well with the pain.	Reflection on effects of thoughts and feelings on behavior	I see that you register that your breathing is not relaxed. Can you give yourself a minute or two to just notice your breathing? Maybe you can find a quiet spot and close your eyes. You could try breathing deeply and slowly a couple of times. Try focusing only on your breath. If you want, you can listen to the instructions to a short mindfulness breathing exercise on the smartphone/website. All the best, Ann.
Values and values-based action	Awareness, planning and evaluation supported by keeping a diary	Today, I plan to [multiple choices possible]: take a walk/work/rest lying down/do household chores/do relaxation exercises/take care of children or others/eat regularly/exercise at a moderate tempo/do my stretching exercises/spend time with family/rest sitting down/spend time with friends/do some shopping/do aerobic exercises/do something just for the pleasure of it.	Reflection on values and values-based behavior based on reports in diaries	I see you have done your stretching exercises today despite reporting a pain level of 6 (scale from 0 to 10; 0=no pain, 10=worst imaginable pain). Can you give yourself a moment to reflect on why this is something you value and choose to do? I would like to ask you to reflect again on your values, if you are willing to, over the next few days. Values are qualities we ourselves think are important and can give us a sense of direction in life. We can ask ourselves questions like: What kind of a person would I like to be in my relations with my family? What can I do today that would get me a bit closer to this ideal? Is this something I am willing to do? Our values are something we can continuously work toward (like being a caring friend), not something we will obtain once and for all. Have a nice weekend, Ann.
Acceptance vs avoidance	Awareness of a spectrum of pain-related thoughts, feelings, and behavior supported by keeping a diary	(1) Right now, I am afraid to be active because of my pain.” (2) Right now, I feel my life is good despite my pain. (3) Right now, I am doing what I want to even if it means increased pain.	Supporting willingness to act in accordance with values despite pain or discouraging thoughts and feelings	(1) I see that today you are not too pleased with your life. Can you give yourself a moment and reflect on what you would want to do today if you were pain free? Is it possible for you to take a small step toward what you want even with your pain? Could you, instead of saying, “I want this, BUT I have pain and therefore can’t” say “I experience pain AND I am taking baby steps toward something valuable to me.” Are you willing to take small steps? (2) Last night you reported a pain level of 8 and that you felt relaxed, grateful, and pleased with the day’s activity level. Can you take a moment to reflect on what kind of self-management strategies you used yesterday? All the best, Ann

The smartphone intervention had the following 4 components:

1. Face-to-face session. The intervention started with a 1-hour individual session between a nurse working on the project and the participant. The session took place in the final week before discharge. Each participant was informed about the intervention and asked about functioning, goals for health-related behavior, and support needs. Values and values-based activities were discussed and the patient received 2 written values-based exercises to take home. The participant was lent a smartphone (HTC TyTN) with a touchscreen and a keyboard. The participants received information (name and qualifications) about their therapist for the intervention (in some cases this was the nurse at the meeting). The nurse attending the face-to-face session summarized the meeting and sent it to the relevant therapist.

2. Web-based diaries. The participant was asked to complete 3 diary entries per day using the smartphone. See Figure 2 for a view of the screen display. The diaries included 16 to 24 questions about the current level and interference of pain, and feelings and thoughts related to avoidance, catastrophizing, and acceptance. They also included questions about planned and previous use of self-management activities and daily values-based and practical activities. Lists of self-management activities (eg, mild exercise, stretching, resting, aerobic exercise, and pleasurable activity) were provided as a reminder. The questions were chosen to support self-monitoring and reflection

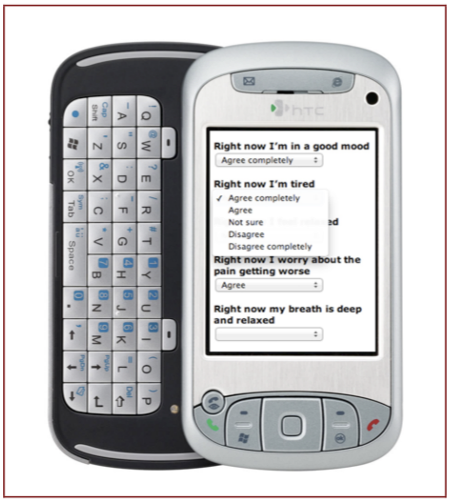
and were formulated in accordance with the experience sampling method principles designed to capture experience in real time without retrospective bias (eg, “Right now I am feeling...”) [45]. See Table 1 for examples of the questions. Participants answered most questions by choosing predefined alternatives or scoring on 5-point Likert scales as shown in Figure 2. All diaries included a comment field giving participants the opportunity to write a short personal message to the therapist. The morning and evening diary entries were sent at fixed hours chosen by each participant. The second diary entry of the day was sent at a time randomly chosen by the Web server, between 11 am and 2 pm. The purpose of including 3 diary entries, including 1 at a randomly chosen time, was to encourage self-monitoring and reflection at different hours and in different situations. At the time scheduled for diary completion, the participant received a Short Message Service (SMS) message with a link to a secure website, where the diary could be opened and questions answered and posted. The participants completed the first diary entry during the face-to-face session, and continued during the final week before discharge with the goal of getting used to the diaries before discharge (a run-in period). After discharge, the diaries were received for 4 weeks. The participant could call a member of the research group (OBK or HE) for technical support. No data were kept on the mobile phone. Up to two automated SMS reminders were sent, if the participant had not responded within 20 to 40 minutes after receiving the SMS signalizing a diary form.

3. Written situational feedback. For 4 weeks after discharge, excluding weekends, participants received daily written feedback from a therapist on the information they had provided in their diaries. The feedback was personalized according to each participant’s situation as reported in the diary. It was written in an empathic style and included repetition of content reported in the diaries, positive reinforcement, reminders of self-management information given at the rehabilitation center, ACT exercises, and reflective questions. The aim was to encourage nonjudgmental awareness of catastrophizing and to stimulate mindfulness and willingness to engage in meaningful activities despite pain or other discouraging intrusions (Table 1). The instructions for the exercises were written directly in the feedback or the participant was referred to exercises available on the mobile phone and/or the website. The feedback was also personalized according to the summary of personal information given at the face-to-face session (eg, family situation and health-related goals) and results on self-reported discrepancy between values and values-based living assessed with the CPVI at the end of the rehabilitation program. The feedback was usually available for the participant within 90 minutes of completing the second diary of the day. If this diary was not submitted, feedback based on information from the most recent submitted diary was sent. When the feedback was available, the participant received an SMS with a link to the website where the feedback could be found. There was no limitation on the length of the feedback, which ranged from a few sentences to a few paragraphs.

The feedback was written by any of 3 of the authors (OBK, TLS, and HE); each participant received signed feedback from the same person throughout the intervention. All therapists had a background in health care sciences (nursing and/or psychology) and had received training in ACT. The feedback protocol was based on ACT for chronic pain [28,34] with a different focus during each of the 4 weeks. For example, in the first week, the focus was on supporting the participant to continue doing the exercises/stretches as recommended at the inpatient program, and during the second week, simple mindfulness exercises were introduced (eg, a few minutes of focused breathing). Once a week, the feedback included an invitation to a values reflection exercise, and every week, questions were included to stimulate reflection on health-related goals. The final feedback comprised a written summary of the registered diary information during the 4-week period. Content from the growing bank of feedback written by all the therapists was used for other participants when appropriate according to the registered information. It took 10 to 15 minutes, on average, to write each piece of feedback. Two members of the group supervised the content of the feedback. They had extensive experience in teaching mindfulness meditation (HE) and supervising CBT/ACT (EAF).

4. Audio files with guided mindfulness exercises. Four audio files with mindfulness exercises (eg, focused breathing) guided by the authors were available on the smartphones.

Figure 2. Smartphone screen display of diary.



Right now I'm in a good mood
Agree completely

Right now I'm tired
✓ Agree completely
Agree
Not sure
Disagree
Disagree completely

Right now I worry about the pain getting worse
Agree

Right now my breath is deep and relaxed

Informational Website With Self-help Pain Management Material

All participants received access to a website with information on self-management strategies for people with chronic pain; not anticipated to have large effect on the study outcomes on its own. It was noninteractive (ie, participants could not register any information or receive feedback). The website included a few written ACT exercises and audio files with mindfulness exercises (as described previously). An example of the written exercises is a behavior analysis aiming to strengthen the ability to observe thought content, feelings, and behavior and the connection between these (adapted from [28]).

Statistical Procedures

Power analyses were based on the level of reported catastrophizing in chronic widespread pain samples [5,20,46,47], a moderate effect size (Cohen's d=0.5), and allowing for attrition commonly seen in studies on Internet interventions [15,48]. A sample size of 70 participants per group was needed to detect a moderate effect size in the primary outcome variable with a 2-sided 5% significance level and 80% power. To investigate differences in demographic variables and baseline characteristics, independent sample t tests, nonparametric tests, and Chi-square tests were used. Data were checked for normal distribution; t tests were used when found suitable for parametric analyses, otherwise nonparametric tests (Mann-Whitney) were applied. The Cohen's d effect sizes were calculated by using the difference between the groups' means divided by the mean standard deviation of both groups. If 1 or 2 items were missing on the GHQ, they were scored as 0 (symptom present less than or as usual). If another instrument included 1 or 2 missing items, the item(s) were replaced with the mean of other items from the participant's instrument. If 2 response alternatives were marked, the healthier option was chosen. Total score was not computed if more than 2 items were missing, and the case was categorized as missing a total score for the instrument. The number of participants included in each analysis is given. In the intention-to-treat analysis, the last observed value was carried forward when data was missing. Five of the participants who withdrew from the smartphone intervention sent in questionnaires at T3 and at the 5-month follow-up (T4). The intention-to-treat analysis included all participants except those who met the exclusion criteria after randomization (n=135). In the analysis of secondary outcomes, only those who completed the interventions were included (n=112). A significance level of P<.05 was used and a tendency toward difference was defined as P<.1. Effect sizes were categorized as small (< 0.5), medium (0.5-0.8), and large (> 0.8) in accordance with Cohen [49].

Results

Participants

A total of 265 women eligible for the study were invited to an informational meeting about the project. Of these, 124 did not attend the meeting or declined to participate. Only 1 was excluded because of a severe psychiatric disorder. One hundred and forty were randomized to the 2 study arms (Figure 1). Five participants met the exclusion criteria after randomization (they were originally submitted for vocational rehabilitation and included in another research project) and 8 discontinued participation before receiving the allocated intervention. In the intervention group, 14 participants did not complete the intervention. Many of those who discontinued participation did so either at the meeting were the allocation information was given or during the intervention's run-in period at the rehabilitation center. The most common reason for withdrawal was finding the participation too stressful in combination with the inpatient program. Another 6 participants discontinued the intervention after discharge from the inpatient program. Demographic data and baseline characteristics of the sample by groups are presented in Table 2. Despite randomization, the groups differed in mean pain level (P=.02) and physical functioning measured by SF-8 (P=.03) at admission to the rehabilitation center. There were no statistically significant differences between the groups at discharge from the rehabilitation center.

Table 2. Participants' characteristics at admission to the inpatient program (T1).

Characteristic		Smartphone intervention (n=69)[a]	Control (n=66)[a]
Age, mean (SD), n		44.59 (11.13), 69	43.80 (11.20), 65
Marital status, n (%)			
	Married or cohabiting	42 (60.9)	45 (68.2)
	Divorced	9 (13.0)	6 (9.1)
	Single	13 (18.8)	10 (15.2)
	Widow	4 (5.8)	2 (3.0)
	Unknown	1 (1.4)	3 (4.5)
Years of education, n (%)			
	10 years (elementary)	13 (18.8)	8 (12.1)
	11-13 years (high school)	19 (27.5)	30 (45.5)
	>13 years (college/university)	30 (43.5)	23 (34.8)
	Unknown	7 (10.1)	5 (7.6)
Employment status, n (%)			
	Working/studying	15 (21.7)	8 (12.1)
	Unemployed	3 (4.3)	1 (1.5)
	On sick leave	27 (39.1)	34 (51.5)
	On disability pension	12 (17.4)	13 (19.7)
	Working/studying part time and part time sick leave	8 (11.6)	5 (7.6)
	Other combination of the above	4 (5.8)	4 (6.1)
	Unknown	0	1 (1.5)
Diagnosed with fibromyalgia, n (%)		55 (80.9)	54 (84.4)
Duration of symptoms (years), mean (SD)		13.11 (8.78)	15.47 (12.09)
Current VAS [b] rating (past couple of days), mean (SD), n			
	Pain	67.08 (17.47), 69	57.85 (21.60), 66
	Fatigue	67.40 (23.73), 69	64.72 (21.02), 66
	Sleep disturbance	57.24 (26.22) 68	55.16 (23.38), 66
Assessments [c] and ranges, mean (SD), n			
	PCS (0-52[d])	21.24 (10.33), 63	20.80 (9.45), 62
	CPAQ (0[d]-120)	56.48 (15.02), 58	53.87 (13.81), 57
	FIQ (0-100[d])	58.75 (16.39), 69	58.58 (16.04), 66
	SF-8, physical (0[d]-100)	31.91 (7.57), 65	34.75 (7.35), 62
	SF-8, mental (0[d]-100)	39.33 (10.49), 65	39.34 (9.61), 62
	GHQ-12 (0-12[d])	3.32 (3.38), 62	3.02 (3.38), 61
	CPVI (0[d]-6)	2.07 (0.95), 64	2.01 (0.73), 61

[a] Patients meeting exclusion criteria after randomization were not included in this analysis.

[b] VAS: visual analog scale, range 0-100.

[c] PCS: pain catastrophizing scale; CPAQ: Chronic Pain Acceptance Questionnaire; FIQ: Fibromyalgia Impact Questionnaire; SF-8: Short-Form Health Survey; GHQ-12: General Health Questionnaire; and CPVI: Chronic Pain Values Inventory.

[d] Values that indicate maximum symptom scores/least health.

Information about fibromyalgia diagnosis was available for 132 participants, and 82.6% of these met the American College of Rheumatology's classification criteria for fibromyalgia. As shown in Tables 3 and 4, no significant group differences were detected at discharge from the rehabilitation center on any of the outcome variables.

Table 3. Means and standard deviations for the primary outcome measure, the pain catastrophizing scale (PCS), at admission to inpatient rehabilitation (T1), at discharge (T2), immediately after intervention (T3), and 5 months after the intervention period (T4).

PCS	T1	T2	T3	T4
ITT and LOCF [a], mean (SD), n				
Intervention	21.24 (10.33), 63	16.06 (10.37), 68	12.32 (9.22), 69	13.59 (9.72), 69
Control	20.80 (9.45), 62	15.33 (9.31), 65	16.07 (9.48), 65	17.43 (11.60), 66
Per protocol, mean (SD), n				
Intervention	20.56 (10.08), 43	14.61 (8.93), 45	9.20 (5.85), 47	10.92 (8.58), 37
Control	20.78 (9.59), 60	15.46 (9.76), 57	15.71 (9.11), 37	18.70 (12.45), 40
PCS score > 24, n (%)				
Intervention	13 (30.2)	7 (15.6)	0 (0)	1 (2.7)
Control	20 (33.3)	10 (17.5)	6 (16.7)	14 (35.0)

[a] ITT: intention-to-treat; LOCF: last observation carried forward.

Table 4. Means and standard deviations for the secondary outcome measures at admission to the inpatient rehabilitation (T1), at discharge (T2), immediately after intervention (T3) and 5 months after the intervention period (T4) for the participants who completed the study.

Secondary outcome measures[a]		T1 mean (SD), n	T2 mean (SD), n	T3 mean (SD), n	T4 mean (SD), n
CPAQ					
	Intervention	56.45 (15.22), 40	62.00 (13.62), 44	72.50 (15.67), 44	71.42 (18.38), 36
	Control	53.94 (13.92), 56	62.21 (10.15), 57	63.55 (13.33), 38	62.47 (14.87), 38
FIQ					
	Intervention	58.46 (17.26), 48	46.38 (16.92), 47	49.12 (19.65), 47	46.45 (19.37), 37
	Control	58.35 (16.18), 64	49.10 (17.32), 62	53.07 (18.68), 39	59.92 (16.46), 40
SF-8, physical					
	Intervention	32.12 (7.74), 45	36.68 (8.42), 40	35.24 (8.74), 46	37.54 (9.44), 37
	Control	34.98 (7.13), 60	35.86 (8.24), 49	36.55 (8.17), 37	34.37 (8.59), 40
SF-8, mental					
	Intervention	39.50 (10.67), 45	45.70 (8.06), 40	46.82 (8.85), 47	44.34 (10.42), 37
	Control	39.09 (9.61), 60	44.83 (9.69), 49	41.01 (9.70), 37	39.78 (10.70), 40
GHQ-12					
	Intervention	3.19 (3.21), 43	1.20 (2.02), 45	1.78 (2.51), 46	1.89 (2.57), 37
	Control	2.97 (3.43), 59	0.63 (1.01), 57	1.86 (2.07), 37	2.85 (3.25), 40
CPVI					
	Intervention	2.05 (0.95), 44	2.47 (0.91), 46	2.95 (0.99), 46	2.62 (0.93), 37
	Control	2.02 (0.74), 59	2.52 (0.68), 54	2.35 (0.91), 38	2.27 (0.83), 40
Pain, VAS					
	Intervention	66.59 (17.58), 48	53.07 (22.20), 47	54.14 (24.06), 47	51.96 (23.76), 37
	Control	57.32 (21.56), 64	52.99 (21.27), 61	50.56 (23.37), 40	58.45 (22.46), 40
Fatigue, VAS					
	Intervention	69.29 (23.98), 48	51.38 (27.75), 47	52.26 (29.18), 47	55.24 (25.73), 37
	Control	64.08 (21.01), 64	50.10 (24.28), 61	53.20 (24.04), 40	65.03 (21.64), 40
Sleep disturbance, VAS					
	Intervention	54.77 (26.99), 47	43.97 (25.77), 47	43.41 (30.60), 47	43.32 (27.88), 37
	Control	54.59 (23.31), 64	48.12 (24.57), 62	48.90 (26.12), 40	57.68 (24.67), 40

[a] CPAQ: Chronic Pain Acceptance Questionnaire; FIQ: Fibromyalgia Impact Questionnaire; SF-8: Short-Form Health Survey; GHQ-12: General Health Questionnaire; CPVI: Chronic Pain Values Inventory; and VAS: visual analog scale.

Within-Group Analysis

Temporal changes within groups and effect sizes within the groups are presented in Tables 5 and 6. Analysis according to the intention-to-treat principles showed a small positive effect on catastrophizing in the intervention group at both assessments and a small negative effect was shown at the 5-month follow-up in the control group. Per-protocol analysis revealed moderate effects on catastrophizing, pain acceptance, and success in living according to values in the intervention group immediately after the follow-up intervention period. The control group did not improve on these variables. The percentage of participants with a total score above 24 on the PCS decreased in the smartphone group, but not in the control group. Increased emotional distress was reported in the control group at 5-month follow-up. Function and symptom impairment, as measured by the FIQ, was increased at both measurements in the control group only. Six months after discharge from the rehabilitation center (5-month follow-up, T4), the improvement in catastrophizing and pain acceptance remained for the intervention group. The changes in success in values-based living were not maintained. However, the control group reported less success in values-based living at the 5-month (T4) follow-up compared to the level at discharge. Pain level was stable in both groups. Fatigue had increased in the control group at the 5-month follow-up and there was a tendency toward more sleep disturbance, which was not seen in the intervention group. Factor analysis of the Norwegian version of the CPAQ revealed some inconsistencies with the 2-factor structure of the scale; 4 items were found to not fit the originally described structure [38]. Because we do not report on the questionnaire’s subscale, we decided to include all questions in our analysis. The results did not differ in a significant way when the 4 items were excluded.

Table 5. Mean differences for the primary outcome measure, the pain catastrophizing scale (PCS) within groups, confidence intervals (CI), and effect sizes (ES).

PCS	Mean difference T2–T3[a] (n)	95% CI T2–T3[a]	Mean difference T2–T4[b] (n)	95% CI T2–T4[b]	ES T2–T3[a]	*P* value[c]	ES T2–T4[b]	*P* value[c]
ITT and LOCF [d]								
Intervention	–3.65 (68)	–5.24 to –2.07	–2.37 (68)	–4.32 to –0.41	0.37	< .001	0.24	.02
Control	0.74 (65)	–0.70 to 2.17	2.30 (65)	0.43-4.16	–0.08	.31	–0.22	.02
Per protocol								
Intervention	–5.09 (44)	–7.00 to –3.18	–2.96 (36)	–5.78 to –0.13	0.69	< .001	0.33	.04
Control	1.67 (34)	–1.06 to 4.40	2.58 (37)	–0.37 to 5.53	–0.18	.22	–0.24	.09

[a] T2: at discharge; T3: immediately after intervention.

[b] T2: at discharge; T4: 5 months after intervention.

[c] *P* values for paired samples *t* tests.

[d] ITT: intention-to-treat; LOCF: last observation carried forward.

Table 6. Mean differences for the secondary outcome measures within groups, confidence intervals (CI), and effect sizes (ES) for the completers.

Secondary outcome measures[a]	Mean difference T2–T3[b] (n)	95% CI T2–T3[b]	Mean difference T2–T4[c] (n)	95% CI T2–T4[c]	ES T2–T3[a]	*P* value[d]	ES T2–T4[c]	*P* value[d]
CPAQ								
Intervention	8.75 (40)	5.96-11.54	7.29 (34)	3.11-11.47	0.58	< .001	0.45	.001
Control	0.69 (36)	–2.90 to 4.29	0.40 (35)	–3.43 to 4.23	0.06	.70	0.03	.83
FIQ								
Intervention	3.10 (46)	–1.01 to 7.20	1.60 (36)	–4.40 to 7.60	–0.17	.14	–0.09	.59
Control	6.61 (38)	2.14-11.09	10.46 (39)	6.43-14.49	–0.36	.005	–0.62	< .001
SF-8, physical								
Intervention	–1.69 (39)	–3.96 to 0.59	0.06 (30)	–3.73 to 3.86	–0.19	.14	0.01	.97
Control	–1.17 (29)	–4.18 to 1.83	–2.41 (32)	–5.17 to 0.36	–0.14	.43	–0.29	.09
SF-8, mental								
Intervention	1.28 (39)	–1.72 to 4.28	–1.51 (30)	–4.93 to 1.92	0.15	.39	–0.17	.38
Control	–3.59 (29)	–7.04 to –0.14	–4.92 (32)	–9.55 to –0.30	–0.37	0.04	–0.50	.04
GHQ-12								
Intervention	0.58 (43)	–0.06 to 1.22	0.80 (35)	–0.42 to 2.02	–0.25	.07	–0.34	.19
Control	1.26 (34)	0.59-1.94	2.38 (37)	1.12-3.63	–0.80	.001	–1.09	< .001
CPVI								
Intervention	0.49 (44)	0.26-0.72	0.15 (36)	–0.11 to 0.42	0.52	< .001	0.16	.25
Control	–0.22 (33)	–0.49 to 0.06	–0.47 (34)	–0.83 to –0.10	–0.28	0.12	–0.63	.01
Pain, VAS								
Intervention	1.11 (46)	–3.94 to 6.16	0.61 (36)	–7.03 to 8.24	–0.05	.66	–0.03	.87
Control	–0.99 (38)	–7.48 to 5.50	5.82 (38)	–1.26 to 12.90	0.04	.76	–0.26	.10
Fatigue, VAS								
Intervention	1.13 (46)	–4.94 to 7.21	7.73 (36)	–2.26 to 17.72	–0.04	.71	–0.29	.13
Control	5.44 (38)	–1.13 to 12.01	12.15 (38)	6.29-18.00	–0.22	.10	–0.51	< .001
Sleep disturbance, VAS								
Intervention	–0.14 (46)	–7.03 to 6.76	2.15 (36)	–7.81-12.12	0.01	.97	–0.08	.66
Control	3.96 (39)	–5.42 to 13.33	7.66 (39)	–0.62 to 15.95	–0.15	.40	–0.30	.07

[a] CPAQ: Chronic Pain Acceptance Questionnaire; FIQ: Fibromyalgia Impact Questionnaire; SF-8: Short-Form Health Survey; GHQ-12: General Health Questionnaire; CPVI: Chronic Pain Values Inventory; and VAS: visual analog scale.

[b] T2: at discharge; T3: immediately after intervention.

[c] T2: at discharge; T4: 5 months after intervention.

[d] *P* values for paired samples *t* tests.

Between-Group Analysis

The between-group effect sizes are shown in Table 7. The intention-to-treat analysis showed a small effect between the groups after the intervention and a tendency (*P*=.05) toward a small effect at 5-month follow-up. The effect size on catastrophizing for completers was large immediately after the intervention period and remained moderate at the 5-month follow-up. Moderate effect sizes were also found for acceptance at both assessments times. There was a moderate effect on values-based living right after the smartphone intervention and a tendency toward moderate effects at 5-month follow-up. A moderate effect on sleep disturbance was found at the 5-month follow-up and a tendency toward moderate effect on fatigue. No effect was found on pain level. A moderate effect was found for functioning and symptom severity measured by the FIQ.

Table 7. Between-group effect sizes (ES) after the smartphone intervention (T3) and at 5-month follow-up (T4).

Outcome measure[a]		ES at T3	*P* value[b]	ES at T4	*P* value[b]
Primary					
	PCS (ITT and LOCF)	0.40	.01	0.36	.05
	PCS (per protocol)	0.87	< .001	0.74	.003
Secondary (per protocol)					
	CPAQ	0.62	.007	0.54	.02
	FIQ	0.21	.35	0.75	.001
	SF-8, physical	–0.15	.64	0.35	.13
	SF-8, mental	0.63	.005	0.43	.06
	GHQ-12	0.03	.56	0.33	.16
	CPVI	0.63	.005	0.40	.08
	Pain, VAS	–0.15	.49	0.28	.22
	Fatigue, VAS	0.04	.87	0.41	.07
	Sleep disturbance, VAS	0.19	.36	0.55	.02

[a] PCS: pain catastrophizing scale; ITT: intention-to-treat; LOCF: last observation carried forward; CPAQ: Chronic Pain Acceptance Questionnaire; FIQ: Fibromyalgia Impact Questionnaire; SF-8: Short-Form Health Survey; GHQ-12: General Health Questionnaire; CPVI: Chronic Pain Values Inventory; and VAS: visual analog scale.

[b] *P* values for independent *t* tests or nonparametric tests.

Withdrawal From Participation

Of the 135 participants, 112 completed the study period (Figure 1). Twenty-one withdrew from the intervention group (30.4%) and 2 withdrew from the control group (3.0%). Because of the small size of the withdrawal group (n=23), group differences with *P*<.2 are described here. The participants who completed the study tended to be younger (mean 43.33, SD 11.18) than the ones who withdrew (mean 48.43, SD 10.06, *P*=.07). There was a tendency toward higher pain level at admission in the group who withdrew (mean 68.79, SD 17.48) than in the group who completed (mean 61.29, SD 20.39, *P*=.15). There was also a tendency toward a higher level of sleep disturbance in the group who withdrew (mean 63.68, SD 23.81) than in the group who completed (mean 54.67, SD 24.81, *P*=.11). A tendency toward a difference was seen in physical functioning as measured by the SF-8; those who withdrew had lower functioning (mean 31.11, SD 7.67) compared to those who completed (mean 33.76, SD 7.50, *P*=.14). At discharge from the rehabilitation center, those who withdrew had more self-reported success in values-based living (mean 2.82, SD 0.91) compared to those who completed (mean 2.50, SD 0.79, *P*=.12).

Response Rates to Assessment Questionnaires

In accordance with the intention-to-treat principle, the response rate for all included participants was 68.1% at T3 (immediately after the smartphone intervention) and 62.2% at T4 (5-month follow-up). There was a higher response rate in the intervention group (75.4%) than in the control group (60.6%) at T3, but the rates were similar at T4 (63.8% and 60.6%, respectively). When only the completers were included in the analysis, more differences in response rates were found. The response rate for the intervention group was 97.9% at T3 and 77.1% at T4. The response rate was 62.5% in the control group at both T3 and T4. The numbers of participants excluded because more than 2 items were missing varied and the number included in each instrument analysis is shown in Table 3.

Comparison of demographic and outcome variables at baseline (T1) between participants who completed the study and returned questionnaires at T3 (n=87) and those who did not return them (n=25) revealed a few differences. Because of the small size of the group who did not return the questionnaires, group differences with *P*<.2 are described here. Those who did return T3 questionnaires had less function impairment and symptom levels at discharge (mean 46.10, SD 17.18) measured by FIQ compared to those who did not return the questionnaire (mean 54.38, SD 15.58, *P*=.04). The same trend was seen in the results of the physical component of the SF-8; those who did return the questionnaires at T3 had better physical functioning (mean 37.20, SD 8.32) at discharge compared to those who did not return them (mean 32.87, SD 7.38, *P*=.04). Those who completed the study, but did not return the questionnaires at T4 (n=35) had lower scores on the CPVI (success scale) (mean 1.78, SD 0.77) at T1 than those returning the questionnaires (mean 2.16, SD 0.84, *P*=.03). At baseline (T1), those who returned T3 questionnaires had higher pain level (mean 62.84, SD 20.02) than those not returning them (mean 55.90, SD 21.17, *P*=.12). There was a tendency toward higher pain level (mean 63.11, SD 20.06) at T1 in those who returned questionnaires at the 5-month follow-up (T4) than those who did not return them (mean 57.29, SD 20.84, *P*=.16). There was also a tendency toward having experienced pain for longer time (mean 17.28 years, SD 13.51) in those who did not return questionnaires at T4 compared with those who did return them (mean 12.77, SD 9.62, *P*=.07). Also, there was a tendency toward worse physical functioning at discharge (mean 34.24, SD 8.54) measured with

SF-8 in those who did not return questionnaires at T4 compared to those who did return them (mean 37.10, SD 8.08, *P*=.14).

Response Rate to the Smartphone Diary Entries and Experience of Participation

The response rate to the diary entries during the 4 weeks after discharge ranged from 27.4% to 95.2%, with a mean of 68.5% and a median of 70.2%. Most (83.3%) participants received 84 entries (4 weeks). A total of 16.7% received additional days of entries to compensate for holidays to ensure 20 days with registration and feedback. Of the 48 participants who completed the study in the smartphone intervention, 43 reported on the experience of participating. Ten (23.3%) participants agreed somewhat that the participation had been experienced as a burden, 9 (20.9%) were neutral in their opinion, 9 (20.9%) disagreed somewhat to the statement, and 15 (34.9%) totally disagreed with the statement that participation was experienced as a burden. Of those who completed the study, 37 (86.0%) agreed somewhat or totally that participation was useful. Three participants (7.0%) were neutral toward this item, and 3 (7.0%) participants disagreed somewhat or totally that participation was useful.

Use of the Informational Website

Of the participants who completed the study in the smartphone intervention, 22 (45.8%) reported never visiting the website. Six (12.5%) visited it once, 8 participants (16.7%) viewed it twice, and 11 (22.9%) viewed it three times or more. One participant did not respond to the question. In the control group, 38 participants who completed the study answered the question. Twelve (18.8%) reported never having visited the website, 5 (7.8%) viewed it once, 9 (14.1%) viewed it twice, and 12 (18.8%) visited 3 or more times.

Discussion

Principal Results

To our knowledge, this is the first study to investigate the efficacy of a smartphone-delivered intervention aiming to reduce catastrophizing and increase functioning in patients with chronic widespread pain. The results from the per-protocol analysis indicate that this intervention with diaries and written personalized feedback reduced catastrophizing and increased acceptance in women with chronic widespread pain and that these effects persisted 5 months after the intervention. At the 5-month follow-up, the control group experienced increased emotional distress compared to the distress at discharge from the inpatient program, whereas the smartphone group did not. The between-group effect size on functioning and symptom level was moderate (0.75) at the 5-month follow-up measured with the FIQ, but no difference was seen in the physical component of the SF-8. One reason for this may be the general nature of the items in the SF-8 compared to the questions in FIQ, possibly making it less sensitive to changes. The results also show a tendency toward increased improvement in values-based living in the intervention group compared to the control group.

When all randomized participants were included in the analysis, the effect size of catastrophizing was small. This may partly be explained by the higher rate of nonresponse in the control group and the method of carrying the last observed value forward resulting in the possibility of a false positive effect for the control group. Scores above 24 on the PCS have been categorized as indicating a high risk for reduced functioning [37]. None of the 7 participants who exceeded this limit before starting the smartphone intervention did so at the end of the intervention. Only one participant was above this criterion again 5 months later. The opposite trend was seen in the control group; an increased number of participants were classified as "catastrophizers."

The intervention was based on CBT, one of the most commonly used models of change in Internet intervention research [50]. We used CBT-related ACT, in which the goal is not to change or reconstruct the content of thoughts, but rather to change how it influences behavior. Behavior change is supported when patients learn to mindfully observe and accept inner experiences and to commit to values-based activity despite challenging thoughts, feelings, or symptoms [28,29,51]. By doing this, the influence of catastrophizing thoughts is expected to be diminished, but by a process other than that described in more traditional CBT, where problematic thoughts are more rationally challenged [28,52]. The reduction in catastrophizing and the increase in acceptance support previous studies that show negative correlations between mindfulness and acceptance, and catastrophizing [37,53,54]. The changes in catastrophizing, acceptance, and functioning in those who completed the study cannot be attributed to changes in levels of pain, or vice versa, because no significant reduction in pain level was found. This is in line with the findings of a recent randomized controlled trial in which fibromyalgia patients who had participated in a 12-week group-based ACT reported more improvement of the condition compared to a waiting-list control group despite no changes in pain level [55]. However, our results differ from that of other previous studies—a small effect size on pain intensity was found in a meta-analysis including 9 randomized trials of acceptance-based interventions [56]. This finding may be explained in part by the fact that the present intervention follows another intervention that had reduced the pain level. The control group showed an increased level of fatigue and a tendency toward an increase in sleep disturbance at the 5-month follow-up. This may indicate that the follow-up intervention might have contributed to the prevention of sleep disturbance. A positive correlation has been found between psychological flexibility and improved sleep quality in people with chronic pain [57].

Strengths and Limitations

To our knowledge, no randomized studies on smartphone interventions based on ACT have been previously published. Our results may support the notion that ACT can be successfully used as a framework for smartphone interventions with mainly written communication. This also supports the results of two recent studies of interventions that provided written ACT-based self-help material and weekly telephone support from a therapist, for 6 and 7 weeks, for patients with chronic pain. Both studies found medium to large effect sizes on pain acceptance [58,59]. However, the present intervention contained many possibly active components and the study design did not allow for any distinction between possible mechanisms and explanations. It is possible that the intervention group benefited from having higher expectations of improvement and from the empathic attention and encouragement from a health care provider [50]. As stated in the guidelines for Internet intervention research, it may still be premature to require demonstration of processes of change in Internet interventions because of the newness of the field [50]. Our results are consistent with the findings of a study that tested the efficacy of a similar mobile phone-delivered intervention with diaries and daily CBT-based feedback for patients with irritable bowel syndrome. The intervention reduced catastrophizing thoughts and the effects were maintained at a 3-month follow-up [18]. Our results are also in line with the results of a follow-up telephone intervention for chronic pain patients designed to support self-monitoring, to give a review of learned self-management techniques, and to provide monthly feedback from a CBT therapist after 11 weeks of group CBT. The intervention was found to reduce pain catastrophizing [13].

Studies on Internet-based interventions and interventions using SMS to support self-management of chronic illness show promising results [60-62]. A review of 14 studies that used SMS to support health behavior change included 6 randomized controlled trials. The duration of the interventions varied from 6 weeks to 1 year and the frequency of communication ranged from many times daily to less than monthly. All but one were effective in supporting positive behavior change, with effect sizes ranging from small to large. However, follow-up data was limited [60]. Reviews of Internet-based interventions for patients with various chronic pain conditions indicate a positive effect on pain, but results on psychological outcomes have been inconsistent [14,15,62]. An important feature of a successful therapeutic relationship is the therapist's ability to respond to what the patient expresses, and tailored or personalized messages have been found to be more effective in supporting behavior change than standardized ones [60]. Our intention was to support a therapeutic relationship with the therapist responding to the expressions made in the diaries and with the goal of sending the individualized feedback as soon as practically possible.

The present study has some limitations. The generalizability of the results is reduced by several factors. Firstly, the intervention group had a withdrawal rate of 30% and this might have resulted in differences in the characteristics of completers between groups. Indeed, there was a trend toward the completers being younger and having less pain, less sleep disturbances, and better function measured with SF-8 at baseline. At admission to the inpatient program (T1), the participants in the smartphone intervention group reported higher pain levels and lower physical functioning compared to the control group. At discharge (T2), this difference was no longer evident. This indicates that participants in the smartphone intervention group improved more on those two variables during the inpatient program compared to the control group. It is possible that this implies some not-assessed differences in the groups' characteristics. Our intention with a run-in period during the final week of the inpatient program was to give the patients a chance to get used to the smartphone diaries before returning home. However, our results may indicate that this might not have been suitable for all participants because several participants withdrew during the run-in phase; it might have been more feasible to give the participants the choice of starting the intervention after discharge from the inpatient program. During the inpatient program, the participants had a busy schedule with activities and may, therefore, have experienced adding the smartphone diaries as stressful. They chose to receive their morning and evening diaries at hours suitable for their schedule at home, which may possibly have been inconvenient while still at the rehabilitation center. High withdrawal rates have been a challenge in SMS-based and Web-based interventions [60]. In a review of 17 trials of Internet self-management interventions for people with chronic pain, the withdrawal rate ranged from 6% to 59% with a median withdrawal rate of 27% [14]. Therapist contact and tailored or personalized messages have been found to correlate with lower withdrawal rates, but as our results show, other factors clearly also play roles. Despite the high withdrawal rate, most experienced the present intervention as useful. In a qualitative study with 7 of our participants, the intervention was described as motivating and supportive [63].

Secondly, the response rate to assessment questionnaires was below 70% at both follow-ups; this affects the generalizability of the results because data cannot be assumed to be missing at random. The response rate was different between the groups, with a lower response rate in the control group immediately after the intervention period. This is commonly experienced in randomized controlled studies [64]. Those who did not return questionnaires after the intervention period (T3) had lower pain levels at baseline (T1) than those who did. Also, those who did not return questionnaires after the smartphone intervention period (T3) had more function and symptom impairment at discharge from the center compared with those who returned those questionnaires. Since all except 1 participant in the smartphone group returned the questionnaires after the intervention (T3) and those who did not respond belonged to the control group, it may be that the level of functional impairment and symptom severity for the control group was, in fact, higher. The 5-month follow-up results could also be affected because there was a tendency toward those not returning the questionnaires reporting less pain at baseline (T1) and better functioning and less symptom severity at discharge (T2). Finally, the generalizability is also affected by the fact that just over half of those eligible to participate were included in the study. We do not know if those who chose to participate differed in any way from those who declined participation. The introduction meeting for the study was scheduled during the second week of the rehabilitation program. For some it may have been too early to consider involvement in a follow-up intervention and others may have used the opportunity to prioritize private time in the tight rehabilitation schedule instead of listening to study information. Moreover, in the stress management part of the rehabilitation program, the patients were encouraged to set limits and say no to requests they felt added more stress to their everyday burden. Patients with high self-efficacy regarding coping after discharge may have been more likely to not attend the informational meeting. Also, because all those who were eligible for the study received a short information letter about the study, some may have found the intervention format unsuitable. In a future study, this kind of intervention might be made more feasible by adding a virtual social support group including fellow participants from the inpatient program. The increase in function impairment and symptom levels in the control group after discharge is not in line with the results of a study on 200 patients with chronic widespread pain or fibromyalgia participating in the same kind of 4-week inpatient program at the same rehabilitation center. The results of the study showed significant improvements in functioning and symptom levels, maintained at both 6- and 12-month follow-ups [9]. However, the samples were not identical because we have excluded men in the present study and those submitted primarily for vocational rehabilitation. Selection bias in our sample may also have had an impact (ie, those with positive long-term effects may have elected not to participate in the study). Nevertheless, the results of the long-term effects of multidimensional pain programs are inconclusive and the need for maintenance support has been clearly stated [65,8,11,12].

This smartphone intervention was developed in 2008 and was delivered using first-generation smartphones. Today, the diary part of the intervention can easily be converted to a smartphone application. Future research might investigate whether automatic feedback could be effectively tailored to diaries and integrated in an application to reduce the investment of human resources used in the presented intervention.

Conclusion

Our results give preliminary support to the efficacy of a smartphone intervention for catastrophizing, acceptance, functioning, and symptom level in women with chronic widespread pain. In addition to subgroup analyses of participants and results on long-term effects, research on practice implications, innovation, and added values for the users are needed.

Acknowledgments

The study is primarily funded by the Research Council of Norway (grant number 182014) (OBK, HE, EE and TLS). We would like to thank the women who were willing to participate in the trial and the staff at Jeløy Kurbad Rehabilitation Centre, especially Marie Utnes Borgaas. We would also like to thank Lena Leren for her assistance with the data material and Thore Egeland for statistical advice. We appreciate the assistance from Lance McCracken during the early phase of the study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT eHealth V1.6.1 [66].

[PDF File (Adobe PDF File), 595KB - jmir_v15i1e5_app1.pdf]

References

1. Croft P, Rigby AS, Boswell R, Schollum J, Silman A. The prevalence of chronic widespread pain in the general population. J Rheumatol 1993 Apr;20(4):710-713. [Medline: 8496870]
2. Clauw DJ, Crofford LJ. Chronic widespread pain and fibromyalgia: what we know, and what we need to know. Best Pract Res Clin Rheumatol 2003 Aug;17(4):685-701. [Medline: 12849719]
3. Shipley M. Chronic widespread pain and fibromyalgia syndrome. Medicine 2010;38(4):202-204. [doi: 10.1016/j.mpmed.2009.12.009]
4. Lindell L, Bergman S, Petersson IF, Jacobsson LT, Herrström P. Prevalence of fibromyalgia and chronic widespread pain. Scand J Prim Health Care 2000 Sep;18(3):149-153. [Medline: 11097099]
5. Cöster L, Kendall S, Gerdle B, Henriksson C, Henriksson KG, Bengtsson A. Chronic widespread musculoskeletal pain - a comparison of those who meet criteria for fibromyalgia and those who do not. Eur J Pain 2008 Jul;12(5):600-610. [doi: 10.1016/j.ejpain.2007.10.001] [Medline: 18024204]
6. Flor H, Turk DC. Chronic Pain: An Integrated Biobehavioral Approach. Seattle, WA: IASP Press; 2011.
7. Goldenberg DL, Burckhardt C, Crofford L. Management of fibromyalgia syndrome. JAMA 2004 Nov 17;292(19):2388-2395. [doi: 10.1001/jama.292.19.2388] [Medline: 15547167]
8. Häuser W, Bernardy K, Arnold B, Offenbächer M, Schiltenwolf M. Efficacy of multicomponent treatment in fibromyalgia syndrome: a meta-analysis of randomized controlled clinical trials. Arthritis Rheum 2009 Feb 15;61(2):216-224 [FREE Full text] [doi: 10.1002/art.24276] [Medline: 19177530]
9. Wigers SH, Finset A. [Rehabilitation of chronic myofascial pain disorders]. Tidsskr Nor Laegeforen 2007 Mar 1;127(5):604-608 [FREE Full text] [Medline: 17332816]
10. Turk DC, Sherman JJ. Treatment of patients with fibromyalgia syndrome. In: Psychological Approaches to Pain Management: A Practitioner's Handbook. New York: The Guilford Press; 2002:390-413.
11. Turk DC, Rudy TE. Neglected topics in the treatment of chronic pain patients--relapse, noncompliance, and adherence enhancement. Pain 1991 Jan;44(1):5-28. [Medline: 2038489]

12. Morley S. Relapse prevention: still neglected after all these years. Pain 2008 Feb;134(3):239-240. [doi: 10.1016/j.pain.2007.12.004] [Medline: 18164134]
13. Naylor MR, Keefe FJ, Brigidi B, Naud S, Helzer JE. Therapeutic Interactive Voice Response for chronic pain reduction and relapse prevention. Pain 2008 Feb;134(3):335-345 [FREE Full text] [doi: 10.1016/j.pain.2007.11.001] [Medline: 18178011]
14. Bender JL, Radhakrishnan A, Diorio C, Englesakis M, Jadad AR. Can pain be managed through the Internet? A systematic review of randomized controlled trials. Pain 2011 Aug;152(8):1740-1750. [doi: 10.1016/j.pain.2011.02.012] [Medline: 21565446]
15. Macea DD, Gajos K, Daglia Calil YA, Fregni F. The efficacy of Web-based cognitive behavioral interventions for chronic pain: a systematic review and meta-analysis. J Pain 2010 Oct;11(10):917-929. [doi: 10.1016/j.jpain.2010.06.005] [Medline: 20650691]
16. Buhrman M, Nilsson-Ihrfeldt E, Jannert M, Ström L, Andersson G. Guided internet-based cognitive behavioural treatment for chronic back pain reduces pain catastrophizing: a randomized controlled trial. J Rehabil Med 2011 May;43(6):500-505 [FREE Full text] [doi: 10.2340/16501977-0805] [Medline: 21533329]
17. Luxton DD, McCann RA, Bush NE, Mishkind MC, Reger GM. mHealth for mental health: Integrating smartphone technology in behavioral healthcare. Prof Psychol-Res Pr 2011;42(6):505-512. [doi: 10.1037/a0024485]
18. Oerlemans S, van Cranenburgh O, Herremans PJ, Spreeuwenberg P, van Dulmen S. Intervening on cognitions and behavior in irritable bowel syndrome: A feasibility trial using PDAs. J Psychosom Res 2011 Mar;70(3):267-277. [doi: 10.1016/j.jpsychores.2010.09.018] [Medline: 21334498]
19. Sorbi MJ, Mak SB, Houtveen JH, Kleiboer AM, van Doornen LJ. Mobile Web-based monitoring and coaching: feasibility in chronic migraine. J Med Internet Res 2007;9(5):e38 [FREE Full text] [doi: 10.2196/jmir.9.5.e38] [Medline: 18166526]
20. Kristjánsdóttir Ó, Fors EA, Eide E, Finset A, van Dulmen S, Wigers SH, et al. Written online situational feedback via mobile phone to support self-management of chronic widespread pain: a usability study of a Web-based intervention. BMC Musculoskelet Disord 2011;12:51 [FREE Full text] [doi: 10.1186/1471-2474-12-51] [Medline: 21352516]
21. Rosser BA, McCullagh P, Davies R, Mountain GA, McCracken L, Eccleston C. Technology-mediated therapy for chronic pain management: the challenges of adapting behavior change interventions for delivery with pervasive communication technology. Telemed J E Health 2011 Apr;17(3):211-216. [doi: 10.1089/tmj.2010.0136] [Medline: 21375412]
22. Rosser BA, Eccleston C. Smartphone applications for pain management. J Telemed Telecare 2011;17(6):308-312. [doi: 10.1258/jtt.2011.101102] [Medline: 21844177]
23. Turk DC, Gatchel RJ. Psychological Approaches to Pain Management: A Practitioner's Handbook. New York: The Guilford Press; 2002.
24. Vlaeyen JW, Linton SJ. Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art. Pain 2000 Apr;85(3):317-332. [Medline: 10781906]
25. Sullivan M, Bishop S, Pivik J. The pain catastrophizing scale: development and validation. Psychological Assessment 1995;7(4):524-532. [doi: 10.1037/1040-3590.7.4.524]
26. Vowles KE, McCracken LM, Eccleston C. Processes of change in treatment for chronic pain: the contributions of pain, acceptance, and catastrophizing. Eur J Pain 2007 Oct;11(7):779-787. [doi: 10.1016/j.ejpain.2006.12.007] [Medline: 17303452]
27. Arnow BA, Blasey CM, Constantino MJ, Robinson R, Hunkeler E, Lee J, et al. Catastrophizing, depression and pain-related disability. Gen Hosp Psychiatry 2011 Apr;33(2):150-156. [doi: 10.1016/j.genhosppsych.2010.12.008] [Medline: 21596208]
28. McCracken LM. Contextual Cognitive-Behavioral Therapy for Chronic Pain. Seattle, WA: IASP Press; 2005.
29. Hayes S, Strosahl K, Wilson K. Acceptance and Commitment Therapy: An Experiential Approach to Behavior Change. New York: The Guilford Press; 2003.
30. McCracken LM. In: McCracken LM, editor. Mindfulness and Acceptance in Behavioral Medicine: Current Theory and Practice (The Context Press Mindfulness and Acceptance Practica Series). Oakland, CA: Context Press; 2011.
31. Wetherell JL, Afari N, Rutledge T, Sorrell JT, Stoddard JA, Petkus AJ, et al. A randomized, controlled trial of acceptance and commitment therapy and cognitive-behavioral therapy for chronic pain. Pain 2011 Sep;152(9):2098-2107. [doi: 10.1016/j.pain.2011.05.016] [Medline: 21683527]
32. Vowles KE, Thompson M. Acceptance and Commitment Therapy for chronic pain. In: McCracken LM, editor. Mindfulness and Acceptance in Behavioral Medicine: Current Theory and Practice (The Context Press Mindfulness and Acceptance Practica Series). Oakland, CA: Context Press; 2011.
33. Bailey KM, Carleton RN, Vlaeyen JW, Asmundson GJ. Treatments addressing pain-related fear and anxiety in patients with chronic musculoskeletal pain: a preliminary review. Cogn Behav Ther 2010 Mar;39(1):46-63. [doi: 10.1080/16506070902980711] [Medline: 19697175]
34. Dahl J, Wilson, KG, Luciano C, Hayes SC. Acceptance and Commitment Therapy for Chronic Pain. Reno, CA: Context Press; 2005.
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. Petty RE, Cacioppo JT. The elaboration likelihood model of persuasion. Adv Exp Soc Psychol 1986;19(1):S0065-S2601. [doi: 10.1016/S0065-2601(08)60214-2]
37. Cassidy EL, Atherton RJ, Robertson N, Walsh DA, Gillett R. Mindfulness, functioning and catastrophizing after multidisciplinary pain management for chronic low back pain. Pain 2012 Mar;153(3):644-650. [doi: 10.1016/j.pain.2011.11.027] [Medline: 22240149]
38. McCracken LM, Vowles KE, Eccleston C. Acceptance of chronic pain: component analysis and a revised assessment method. Pain 2004 Jan;107(1-2):159-166. [Medline: 14715402]
39. Goldberg DP, Gater R, Sartorius N, Ustun TB, Piccinelli M, Gureje O, et al. The validity of two versions of the GHQ in the WHO study of mental illness in general health care. Psychol Med 1997 Jan;27(1):191-197. [Medline: 9122299]
40. McCracken LM, Yang SY. The role of values in a contextual cognitive-behavioral approach to chronic pain. Pain 2006 Jul;123(1-2):137-145. [doi: 10.1016/j.pain.2006.02.021] [Medline: 16564627]
41. Burckhardt CS, Clark SR, Bennett RM. The fibromyalgia impact questionnaire: development and validation. J Rheumatol 1991 May;18(5):728-733. [Medline: 1865419]
42. Saris-Baglama RN, Dewey CJ, Chisholm GB, Plumb E, King J, Rasicot P, et al. QualityMetric Outcomes Scoring Software 4.5. User's Guide. Lincoln, RI: Quality Metric Incorporated; 2011.
43. Simons DG, Travell JG, Simons LS. Travell & Simons' Myofascial Pain and Dysfunction: The Trigger Point Manual Volume 1. Baltimore, MD: Williams & Wilkins; 1999.
44. Travell, Simons DG. Travell & Simons' Myofascial Pain and Dysfunction: The Trigger Point Manual Volume 2. Baltimore, MD: Williams & Wilkins; 1992.
45. Scollon C, Kim-Prieto C, Diener E. Experience Sampling: Promises and Pitfalls, Strengths and Weaknesses. Journal of Happiness Studies 2003;4(1):5-34. [doi: 10.1023/A:1023605205115]
46. Van Damme S, Crombez G, Bijttebier P, Goubert L, Van Houdenhove B. A confirmatory factor analysis of the Pain Catastrophizing Scale: invariant factor structure across clinical and non-clinical populations. Pain 2002 Apr;96(3):319-324. [Medline: 11973004]
47. Severeijns R, Vlaeyen JW, van den Hout MA, Picavet HS. Pain catastrophizing is associated with health indices in musculoskeletal pain: a cross-sectional study in the Dutch community. Health Psychol 2004 Jan;23(1):49-57. [doi: 10.1037/0278-6133.23.1.49] [Medline: 14756603]
48. Andersson G. Using the Internet to provide cognitive behaviour therapy. Behav Res Ther 2009 Mar;47(3):175-180. [doi: 10.1016/j.brat.2009.01.010] [Medline: 19230862]
49. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Hillsdale, NJ: L Erlbaum Associates; 1988.
50. Proudfoot J, Klein B, Barak A, Carlbring P, Cuijpers P, Lange A, et al. Establishing guidelines for executing and reporting internet intervention research. Cognitive Behaviour Therapy 2011;40(2):82-97. [doi: 10.1080/16506073.2011.573807]
51. Hayes SC, Luoma JB, Bond FW, Masuda A, Lillis J. Acceptance and commitment therapy: model, processes and outcomes. Behav Res Ther 2006 Jan;44(1):1-25. [doi: 10.1016/j.brat.2005.06.006] [Medline: 16300724]
52. Winterowd C, Beck AT, Gruener D. Cognitive Therapy with Chronic Pain Patients. New York: Springer; 2003.
53. Schütze R, Rees C, Preece M, Schütze M. Low mindfulness predicts pain catastrophizing in a fear-avoidance model of chronic pain. Pain 2010 Jan;148(1):120-127. [doi: 10.1016/j.pain.2009.10.030] [Medline: 19944534]
54. Vowles KE, McCracken LM, Eccleston C. Patient functioning and catastrophizing in chronic pain: the mediating effects of acceptance. Health Psychol 2008 Mar;27(2 Suppl):S136-S143. [doi: 10.1037/0278-6133.27.2(Suppl.).S136] [Medline: 18377155]
55. Jensen KB, Kosek E, Wicksell R, Kemani M, Olsson G, Merle JV, et al. Cognitive Behavioral Therapy increases pain-evoked activation of the prefrontal cortex in patients with fibromyalgia. Pain 2012 Jul;153(7):1495-1503. [doi: 10.1016/j.pain.2012.04.010] [Medline: 22617632]
56. Veehof MM, Oskam MJ, Schreurs KM, Bohlmeijer ET. Acceptance-based interventions for the treatment of chronic pain: a systematic review and meta-analysis. Pain 2011 Mar;152(3):533-542. [doi: 10.1016/j.pain.2010.11.002] [Medline: 21251756]
57. McCracken LM, Williams JL, Tang NK. Psychological flexibility may reduce insomnia in persons with chronic pain: a preliminary retrospective study. Pain Med 2011 Jun;12(6):904-912. [doi: 10.1111/j.1526-4637.2011.01115.x] [Medline: 21539701]
58. Johnston M, Foster M, Shennan J, Starkey NJ, Johnson A. The effectiveness of an Acceptance and Commitment Therapy self-help intervention for chronic pain. Clin J Pain 2010 Jun;26(5):393-402. [doi: 10.1097/AJP.0b013e3181cf59ce] [Medline: 20473046]
59. Thorsell J, Finnes A, Dahl J, Lundgren T, Gybrant M, Gordh T, et al. A comparative study of 2 manual-based self-help interventions, acceptance and commitment therapy and applied relaxation, for persons with chronic pain. Clin J Pain 2011 Oct;27(8):716-723. [doi: 10.1097/AJP.0b013e318219a933] [Medline: 21540740]
60. Fjeldsoe BS, Marshall AL, Miller YD. Behavior change interventions delivered by mobile telephone short-message service. Am J Prev Med 2009 Feb;36(2):165-173. [doi: 10.1016/j.amepre.2008.09.040] [Medline: 19135907]
61. Wei J, Hollin I, Kachnowski S. A review of the use of mobile phone text messaging in clinical and healthy behaviour interventions. J Telemed Telecare 2011;17(1):41-48. [doi: 10.1258/jtt.2010.100322] [Medline: 21097565]

62. Cuijpers P, van Straten A, Andersson G. Internet-administered cognitive behavior therapy for health problems: a systematic review. J Behav Med 2008 Apr;31(2):169-177 [FREE Full text] [doi: 10.1007/s10865-007-9144-1] [Medline: 18165893]
63. Jelin E, Granum V, Eide H. Experiences of a web-based nursing intervention--interviews with women with chronic musculoskeletal pain. Pain Manag Nurs 2012 Mar;13(1):2-10. [doi: 10.1016/j.pmn.2011.08.008] [Medline: 22341135]
64. Streiner DL. Missing data and the trouble with LOCF. Evid Based Ment Health 2008 Feb;11(1):3-5. [doi: 10.1136/ebmh.11.1.3-a] [Medline: 18223040]
65. Karjalainen K, Malmivaara A, van Tulder M, Roine R, Jauhiainen M, Hurri H, et al. Multidisciplinary rehabilitation for fibromyalgia and musculoskeletal pain in working age adults. Cochrane Database Syst Rev 2000(2):CD001984. [doi: 10.1002/14651858.CD001984] [Medline: 10796458]
66. 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

ACT: Acceptance and Commitment Therapy
CBT: cognitive behavioral therapy
CPAQ: Chronic Pain Acceptance Questionnaire
CPVI: Chronic Pain Values Inventory
ES: effect sizes
FIQ: Fibromyalgia Impact Questionnaire
GHQ: General Health Questionnaire
LOCF: last observation carried forward
PCS: Pain Catastrophizing Scale
SD: standard deviations
SDT: self-determination theory
SF-8: Short-Form Health Survey
SMS: Short Message Service
VAS: visual analog scale

Edited by G Eysenbach; submitted 30.06.12; peer-reviewed by J Kelly, J Mooney, S Langrial; comments to author 25.09.12; accepted 04.10.12; published 07.01.13.

Please cite as:
Kristjánsdóttir ÓB, Fors EA, Eide E, Finset A, Stensrud TL, van Dulmen S, Wigers SH, Eide H
A Smartphone-Based Intervention With Diaries and Therapist-Feedback to Reduce Catastrophizing and Increase Functioning in Women With Chronic Widespread Pain: Randomized Controlled Trial
J Med Internet Res 2013;15(1):e5
URL: http://www.jmir.org/2013/1/e5/
doi:10.2196/jmir.2249
PMID:23291270

Original Paper

Patient Attitudes Toward Mobile Phone-Based Health Monitoring: Questionnaire Study Among Kidney Transplant Recipients

John William McGillicuddy[1], MD; Ana Katherine Weiland[2], BS; Ronja Maximiliane Frenzel[2], BS; Martina Mueller[2], RN, PhD; Brenda Marie Brunner-Jackson[2], MPH; David James Taber[1], PharmD; Prabhakar Kalyanpur Baliga[1], MD; Frank Anton Treiber[2,3], PhD

[1]Division of Transplant Surgery, Department of Surgery, Medical University of South Carolina, Charleston, SC, United States
[2]College of Nursing, Medical University of South Carolina, Charleston, SC, United States
[3]College of Medicine, Medical University of South Carolina, Charleston, SC, United States

Corresponding Author:
John William McGillicuddy, MD
Division of Transplant Surgery
Department of Surgery
Medical University of South Carolina
96 Jonathan Lucas St.
CSB 409/ MSC 611
Charleston, SC, 29425-6110
United States
Phone: 1 843 792 3368
Fax: 1 843 792 8596
Email: mcgillij@musc.edu

Abstract

Background: Mobile phone based remote monitoring of medication adherence and physiological parameters has the potential of improving long-term graft outcomes in the recipients of kidney transplants. This technology is promising as it is relatively inexpensive, can include intuitive software and may offer the ability to conduct close patient monitoring in a non-intrusive manner. This includes the optimal management of comorbidities such as hypertension and diabetes. There is, however, a lack of data assessing the attitudes of renal transplant recipients toward this technology, especially among ethnic minorities.

Objective: To assess the attitudes of renal transplant recipients toward mobile phone based remote monitoring and management of their medical regimen; and to identify demographic or clinical characteristics that impact on this attitude.

Methods: After a 10 minute demonstration of a prototype mobile phone based monitoring system, a 10 item questionnaire regarding attitude toward remote monitoring and the technology was administered to the participants, along with the 10 item Perceived Stress Scale and the 7 item Morisky Medication Adherence Scale.

Results: Between February and April 2012, a total of 99 renal transplant recipients were identified and agreed to participate in the survey. The results of the survey indicate that while 90% (87/97) of respondents own a mobile phone, only 7% (7/98) had any prior knowledge of mobile phone based remote monitoring. Despite this, the majority of respondents, 79% (78/99), reported a positive attitude toward the use of a prototype system if it came at no cost to themselves. Blacks were more likely than whites to own smartphones (43.1%, 28/65 vs 20.6%, 7/34; *P*=.03) and held a more positive attitude toward free use of the prototype system than whites (4.25±0.88 vs 3.76±1.07; *P*=.02).

Conclusions: The data demonstrates that kidney transplant recipients have a positive overall attitude toward mobile phone based health technology (mHealth). Additionally, the data demonstrates that most kidney transplant recipients own and are comfortable using mobile phones and that many of these patients already own and use smart mobile phones. The respondents felt that mHealth offers an opportunity for improved self-efficacy and improved provider driven medical management. Respondents were comfortable with the idea of being monitored using mobile technology and are confident that their privacy can be protected. The small subset of kidney transplant recipients who are less interested in mHealth may be less technologically adept as reflected by their lower mobile phone ownership rates. As a whole, kidney transplant recipients are receptive to the technology and believe in its utility.

(J Med Internet Res 2013;15(1):e6) doi:10.2196/jmir.2284

KEYWORDS
cellular phone; attitude; kidney transplantation; telemedicine; mobile phone; smartphone; mhealth

Introduction

Nearly 400,000 people living in the United States suffer with end stage renal disease, of these; approximately 93,000 are awaiting kidney transplantation [1,2]. Kidney transplantation is the preferred mode of treatment for end stage renal disease as it offers superior quality of life and improved life expectancy compared to chronic dialysis [3-6]. Outcomes after kidney transplantation are negatively impacted by poor medication adherence and suboptimal control of common comorbid medical conditions such as hypertension and diabetes [7-10]. Black patients suffer disproportionately with end stage renal disease and represent the vast majority of patients on dialysis in South Carolina [1]. Black kidney transplant recipients suffer a poorer graft survival than white kidney transplant recipients [11-17]. The reasons for this may include poorer medication adherence [11,18], heightened immunological response [19,20] and a higher prevalence of comorbid illnesses [1,21]. The development of effective, efficient and non-intrusive approaches to aid kidney transplant recipients self-management and monitoring is critical to success as limited healthcare provider resources are increasingly taxed by growing demand.

Recent studies have suggested that remote monitoring via mobile health technology (mHealth) is an effective and sustainable strategy for facilitating patient provider communication, improving health outcomes, increasing adherence to medical regimens and reducing costs in some chronic illnesses [22-30]. Mobile phone based monitoring is an attractive option due to their ubiquity, connectivity, computational power, portability and relatively low cost [23,25,31-32]. A critical component to the success of any mHealth system is the willingness and ability of the target population to adopt and effectively utilize the technology. Previous studies have investigated the attitudes of different patient populations in regards to mobile phone based remote monitoring in other disease states [33-36]. Kidney transplant recipients are a unique population due to the high complexity of their medical regimens, the critical importance of strict medication adherence, the near universal presence of significant comorbid medical illness and the geographic distance that often separates them from their transplant center [7,9,37-38]. Kidney transplant recipients also tend to be relatively aged, a factor that may lead this population to be less willing or able to successfully utilize advanced technologies such as smart mobile phones or Bluetooth enabled medical devices.

The aims of this study were twofold. First, to assess the attitudes of a racially diverse sample of kidney transplant recipients on the use of a mHealth remote monitoring system, particularly to enhance medication adherence and blood pressure control. Second, to investigate the whether demographics, prior technology utilization, stress levels and self-reported medication adherence impact on the attitudes toward mHealth.

Methods

Participants and Recruitment

Study participants were recruited from the Kidney Transplant Clinic at the Medical University of South Carolina (MUSC), Charleston. Eligible patients were those who had previously received a kidney transplant, were over 18 years of age and spoke English. Between February and April 2012, 103 patients were approached during their usual post-transplant clinic visit, either by their clinical coordinator, clinic nurse or physician. Patients were asked if they were willing to speak to the study coordinator regarding participating in the survey, 99 agreed to participate. Those that declined did so either for lack of time or lack of interest. The demographic and transplant related clinical characteristics of the survey participants are summarized in Table 1. The study was approved by the MUSC institutional review board.

Study Setting and Design

Patients were approached about completing a survey that included questions on their attitudes toward remote monitoring, mobile phones, electronic medication monitors and electronic home blood pressure monitors. Also evaluated were perceived levels of general stress and medication adherence (Table 1). Patients were individually shown to a private clinic space, accompanied by their informal caregivers (if present), where they were provided a description and demonstration of a prototype mHealth system with a presentation of the specific steps that were required to utilize the device (Multimedia Appendix 1).

The prototype system included a smartphone (Motorola Droid X), a wireless (GSM-enabled) medication tray (Maya MedMinder) and a wireless (Bluetooth-enabled) blood pressure monitor (Fora D15b). Patients were required to use the medication tray for all medication dispensing and to measure their resting blood pressure and pulse every third day at both morning and evening. The smartphone automatically transfers the blood pressure and pulse data to computer servers for later analysis. The medication tray is fully programmable and capable of delivering reminders in the form of light, tone, text message or phone call. Adherence is tracked in real time and can trigger the delivery of motivational or positive reinforcement messages to the patient via text, email or phone. A summary of the adherence over time can be generated and delivered via email to the patient and the treating clinician. Blood pressure and pulse readings outside predetermined safe parameters could generate automated alert messages that would be sent to both patient and physician. Participants were informed that a clinical coordinator would contact them in the event that alerts were generated.

Medication Adherence Scale

Medication adherence was evaluated using a 7 item modified Morisky Medication Adherence Scale with an internal consistency of 0.82 and a sensitivity and specificity of 91% and

50% respectively [39]. Further description and psychometric data on the Morisky scale are described in detail elsewhere [40]. The modified Morisky scale yields a score in the range of 0 to 7 with higher scores reflecting higher adherence to medication. Scores can be categorized into high (=7), moderate (≥6 but <7) and poor (<6) adherence levels based on its criterion validity with blood pressure control among hypertensive patients [39].

Perceived Stress Scale

Generalized perceived stress was measured using the Perceived Stress Scale in which each of the 10 items is answered using a 5 point Likert item ranging from ‘never’ to ‘very often’. The scale has established psychometric properties including internal consistency of 0.85 [41]. Internal consistency for the present study was 0.85.

mHealth Related Survey

The respondents’ awareness of, and attitudes toward, mHealth and telemedicine based remote monitoring were evaluated using a 10 item survey (Table 2), in which 9 of the 10 items were answered using a 5 point Likert item ranging from ‘strongly disagree’ to ‘strongly agree’. The items were adapted from prior studies assessing patients attitudes toward mobile phone based remote monitoring for chronic illness [34,36,42]. Items included questions on their perceptions of remote monitoring and their comfort with using mobile phones. Cronbach’s alpha internal consistency coefficient was 0.92 for the 9 items. The 10th item queried their *a priori* awareness of health related remote monitoring technology.

Our sample was analyzed in two ways: by race, blacks versus whites (Table 3) and by attitude toward use of mHealth. Participants who answered either ‘agree’ or ‘strongly agree’ to the question whether they would use the mHealth system if it were free were compared to the participants who chose either ‘disagree’, ‘strongly disagree’ or ‘neutral’ (Table 4). Means and standard deviations for continuous variables and frequency distributions for categorical variables were used to describe the characteristics of the total sample and the racial and mHealth attitude groups. Clinical and demographic features were compared for racial and attitude groups using the pooled t test for continuous variables and the X^2 test/Fisher exact test for categorical variables.

Results

Demographic and Clinical Characteristics

The results indicate that 85% (83/98) of respondents are presently using devices at home to monitor either their blood pressure or blood sugar. Nearly two thirds of the respondents, 63% (62/99), were using medication dispensing devices (ie. standard non-signaling pillboxes). Respondents reported slightly lower than average levels of perceived stress (Table 1) that were not remarkably higher than the general population (10.0 ± 6.5 out of 40) [43]. Self-reported adherence was also reported to be moderately high (6.1 out of 7) [39]. There were no racial differences in perceived stress (*P*=.13), but there was a difference in self-reported adherence (*P*=.03).

Mobile Phone Utilization

Nearly all respondents (90%, 87/97) indicated that they already own a mobile phone and were comfortable utilizing this technology. Over a third of the patients (35%, 35/99) reported that they own a smartphone and over half (52%, 51/99) reported that they had access to a working smartphone in the household other than their own. Most of the patients (82%, 81/99), indicated that there was someone in the household who could assist them with using a mobile phone if needed.

The survey results also indicated that most respondents had a familiarity with using mobile phones for reasons other than phone calls. Amongst these, over half (61%, 60/99) used a mobile phone to send or receive text messages, 38% (38/99) to browse the web, 35% (35/99) to send or receive email and 34% (34/ 99) reported using a mobile phone to download a ringtone or a mobile application.

Table 1. Demographic and transplant related clinical characteristics of survey participants.

Variable	Mean (± SD) or Proportion
Age	53.1 ± 13.4 (Median 52)
Gender (Male)	65% (64/98)
Race (Black)	66% (65/99)
Marital status (Married)	64% (63/98)
Education level (≤High School)	38% (38/99)
Employment (Part or Full Time)	22% (21/98)
Annual income (<$30,000)	57% (44/77)
Months since kidney transplant	29.2 ± 54.5
More than one transplant	12% (12/99)
Perceived Stress Scale score	10.9 ± 6.5
Morisky Scale score	6.1 ± 1.1

Attitudes and Willingness to Use mHealth Technology

Only 7% (7/98) of participants had any prior knowledge of mHealth remote monitoring technology before being surveyed (Table 2). However, most of the participants felt that mobile technology would be helpful in reminding them to follow their doctor's directions (81%, 80/99). The majority also felt that the technology would allow their doctor to make more rapid adjustments to their medication regimen (84%, 83/99). Furthermore, most of the participants, 79% (78/99), indicated that if they were provided the mHealth system at no cost and instructed on its use that they would use it as directed by their health care provider. The addition of free technical support did not significantly increase their willingness to use the devices. On the matter of health information, 80% (79/99) indicated that they were comfortable with a health care provider monitoring their health information using remote monitoring technology and 76% (75/99) felt confident that their privacy could be adequately protected. Almost all participants (95%, 94/99) felt that it was important to follow their doctor's directions and 87% (86/99) thought that remote monitoring technology would help them effectively communicate with their health care providers about their medical conditions. When asked about how they would prefer to receive instructions from their health care providers, most respondents preferred that communication be done via phone call, with voicemail being the most common second choice. Text messaging was the third most common choice, with only a small fraction of patients indicating that they were interested in receiving instruction primarily via live video conferencing.

Table 2. Responses to mHealth related survey.

Survey Items	Mean (± SD) or Percentage
Heard of tele-health (yes)	7% (7/98)
Would use mHealth devices if free	4.08 ± 0.98
If someone available to answer questions likely to use devices as directed	4.18 ± 0.92
Comfortable having health monitored remotely by doctor/nurses using mHealth technologies	4.16 ± 0.89
Comfortable using cell phone	4.30 ± 0.80
Mobile technology will help remind me to follow doctor's directions	4.14 ± 0.89
Mobile technology will allow doctor to make medication changes quicker	4.22 ± 0.79
Confident privacy protected when using mHealth devices	4.07 ± 0.92
Important to follow doctor's directions	4.56 ± 0.72
Confident mHealth technology can effectively communicate my medical condition to my doctor	4.24 ± 0.81

Racial Differences

As shown in Table 3, when compared to whites, blacks were younger (*P*=.04), more likely to report hypertension as an etiology of their renal failure (*P*=.05), more likely to be a first time recipient of a kidney transplant (*P*=.02) and more likely to live nearer to the transplant center (*P*<.001). While mobile phone ownership did not differ significantly between races, blacks were more likely than whites to own a smartphone (*P*=.03). Although it did not reach statistical significance, there was a trend toward more mobile phone based internet usage among blacks (*P*=.08). Perceived stress levels did not vary by race (*P*=.13), but there was a significant difference in medication adherence with blacks reporting slightly poorer adherence (*P*=.03). Blacks had a slightly more positive attitude toward mHealth than whites as gauged by their willingness to use the technology if it came at no cost to them (*P*=.02). There were no significant differences between blacks and whites in the level of education (*P*=.62), annual income (*P*=.16) or employment status (*P*=.46).

Table 3. Race comparison on transplant related characteristics, stress exposure, medication adherence and cell phone ownership.

		Black	White	Degrees of Freedom	*P*
Age		51 ± 13.1	56.9 ±13.4	96	.04
Gender (Male)		65%	65%	1	.99
Primary cause for kidney failure				6	.05
	HTN[a] (alone or + other)	39%	21%		
	Diabetes (alone or + other)	22%	18%		
	Diabetes + HTN[a]	15%	9%		
	Other	20%	35%		
	Not sure	5%	18%		
More than one transplant (yes)		6%	24%	1	.02
Travel time to transplant center (< 2 hours)		77%	33%	3	<.001
Own mobile phone		86%	97%	1	.16
Own smartphone		43%	21%	1	.03
Would use mHealth devices if free		4.3 ± .88	3.8 ± 1.1	3	.02
Perceived Stress Scale score		11.7 ± 6.8	9.5 ± 5.7	97	.13
Morisky Scale score		5.9 ± 1.2	6.5 ± 0.9	97	.03

[a] Hypertension

Characteristics of Respondents with Positive Attitude Toward mHealth

Respondents who answered either ‘agree’ or ’strongly agree’ to the query as to whether they would use the mHealth system, as demonstrated, if it were free, were more likely to be employed (*P*=.04), the recipient of their first transplant (*P*=.02), already using a medication tray at home (*P*=.04) and the owner of a working mobile phone (*P*=.04). These respondents were also more likely to own a smartphone (*P*=.01) and to have used a mobile phone to text (*P*=.02), email (*P*=.01), browse the internet (*P*=.002) or download an application or ringtone (*P*=.03). As can be seen in Table 4, they also reported higher levels of perceived stress (*P*=.01).

Table 4. Comparisons of patients who do versus do not favor use of mHealth devices.

		Agree	Do Not Agree	Degrees of Freedom	*P*
Age		52 ± 13.1 (Median 51)	57 ± 14.3 (Median 62)	96	.13
Race				1	.01
	Black	86%	14%		
	White	65%	35%		
Socioeconomic status					
	Employment (Part or Full Time)	27%	5%	1	.04
	Education level (≤High School)	37%	43%	2	.89
	Annual income (<$30,000)	45%	43%	1	.54
More than one kidney transplant		8%	29%	1	.02
Use medication dispensing device at home		68%	43%	1	.04
Own cell phone		93%	76%	1	.04
Own smartphone		42%	10%	1	.005
Perceived Stress Scale score		11.7 ± 6.5	7.7 ± 5.4	97	.01
Morisky Scale score		6.0 ± 1.2	6.6 ± 0.7	97	.04

Discussion

Recent literature demonstrates that mHealth technology can have a positive impact on the quality of life, self-efficacy and the ability to monitor biochemical or physiologic markers of disease control across a wide array of illnesses [44]. While the evidence is mixed as to the cost effectiveness of mHealth technology at present [45], it seems reasonable to hypothesize that it will become cost effective, as the demand increases, cost of the technology decreases and the long-term health benefits are realized. Furthermore, as penetrance of the smartphone technology increases amongst consumers, it seems likely that there will be an increasing demand for this type of health care delivery from patients.

Principal Results

With 90% (87/97) of respondents owning a mobile phone and 35% (35/99) owning a smartphone, this population of kidney transplant recipients closely mirrors the adult American population [46]. This finding is mildly surprising given South Carolina's historically low household income and underscores the near ubiquitous use of mobile phone technology. Interestingly, in our cohort, while there were no racial differences in overall mobile phone ownership, blacks were significantly more likely to own a smartphone. This reflects national figures that show higher rates of smartphone adoption among racial minority groups [47]. Blacks are significantly more likely than whites to use a wireless device to access the internet [46]. As early adopters and high utilizers, blacks may be uniquely positioned to benefit from improving mHealth technology. The penetrance of mobile phone technology, particularly the rapid ascension of the smartphone [47], bodes well for the continued expansion of the mobile phone's role in health care delivery.

Few of the kidney transplant recipients had any knowledge of mHealth technology prior to being surveyed. Despite that, the vast majority was receptive to utilizing such technology if the devices were provided at no cost. Respondents felt that having the technology would help them follow their medical regimen and improve communication with their healthcare providers, particularly with regards to the efficiency of regulating or changing their medical regime as the need arises. These findings are consistent with other studies that have evaluated the attitudes toward mHealth technology among patients with various chronic illnesses, including essential hypertension, diabetes and congestive heart failure [27,34,36].

Although there was a high receptivity toward using mHealth technology, there was a cohort of respondents who indicated a less than positive attitude. That these respondents were less likely to own a mobile phone and far less likely to own a smartphone might reflect a lower level of comfort with technology. This potential barrier to use of an mHealth system could be addressed both by making the system easier to use and by providing some skilled assistance and training. The fact that these same respondents were more likely to have had a prior kidney transplant and were less likely to be using a medication tray at home, might indicate a higher comfort level with immunosuppression medications or a lower perceived importance of medication adherence. That this cohort self-reports higher adherence with medications and lower levels of perceived stress may, in part, explain their diminished interest in the technology. Unsurprisingly, this same cohort of patients was less likely to respond positively to the remaining questions regarding comfort with being monitored, comfort with mobile phone technology, privacy protection and various aspects of utility.

Limitations

These findings must be evaluated within the context of several limitations of the study. First, that all respondents were recruited from a single transplant center, which may call into question the generalizability of the findings. However, it should be noted that this center is the sole transplant service provider for the State of South Carolina and has a catchment population of over 4.6 million persons. Second, those that chose to participate might be predisposed to a positive attitude toward mHealth and thereby introduce a positive bias. The participation of nearly everyone who was approached however, suggests that a significant bias toward mHealth is unlikely. Third, it cannot be assumed that the respondents' purported interest in mHealth will translate into actual use. Anecdotally, as we have begun to enroll kidney transplant recipients in a mHealth medication adherence trial, as a proof of concept research based on this work, we have experienced high participation and utilization rates.

Conclusions

This is the first study assessing the attitudes of transplant recipients with this technology and the data demonstrates that there is a positive overall attitude towards mHealth technology. Additionally, the data demonstrates that most kidney transplant recipients already own and are comfortable using mobile phones and that many of these participants already own and use smart mobile phones. Results indicate that the participants feel that mHealth offers an opportunity for improved self-efficacy and improved provider driven medical management. Participants are also comfortable with the idea of being monitored using mobile technology and are confident that their privacy can be protected. As a whole, kidney transplant recipients are receptive to the technology and believe in its utility. Further research in this area should include patient centered evaluations of usability and usefulness as well as proof of concept trials to identify areas of concern.

Acknowledgments

This publication was supported by the South Carolina Clinical and Translational Research (SCTR) Institute, with an academic home at the Medical University of South Carolina, National Institutes of Health, Grant number UL1 RR029882, KL2 RR029880, UL1 TR000062 and KL2 TR000060. The contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health or the National Center for Research Resources.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Patient demo video.

[MP4 File (MP4 Video), 71MB - jmir_v15i1e6_app1.mp4]

References

1. U.S. Renal Data System.: National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases; 2011. Annual data report: Atlas of chronic kidney disease and end-stage renal disease in the United States URL: http://www.usrds.org/2011/view/ [accessed 2012-12-29] [WebCite Cache ID 6DHuy2NCW]
2. OPTN: Organ Procurement and Transplant Network.: U.S. Department of Health and Human Services; 2011. URL: http://optn.transplant.hrsa.gov/ [accessed 2012-07-23] [WebCite Cache ID 69NPbLSaB]
3. Manninen DL, Evans RW, Dugan MK. Work disability, functional limitations, and the health status of kidney transplantation recipients posttransplant. Clin Transpl 1991:193-203. [Medline: 1820116]
4. Laupacis A, Keown P, Pus N, Krueger H, Ferguson B, Wong C, et al. A study of the quality of life and cost-utility of renal transplantation. Kidney Int 1996 Jul;50(1):235-242. [Medline: 8807593]
5. Wolfe RA, Ashby VB, Milford EL, Ojo AO, Ettenger RE, Agodoa LY, et al. Comparison of mortality in all patients on dialysis, patients on dialysis awaiting transplantation, and recipients of a first cadaveric transplant. N Engl J Med 1999 Dec 2;341(23):1725-1730. [doi: 10.1056/NEJM199912023412303] [Medline: 10580071]
6. Neipp M, Karavul B, Jackobs S, Meyer zu Vilsendorf A, Richter N, Becker T, et al. Quality of life in adult transplant recipients more than 15 years after kidney transplantation. Transplantation 2006 Jun 27;81(12):1640-1644. [doi: 10.1097/01.tp.0000226070.74443.fb] [Medline: 16794528]
7. De Geest S, Borgermans L, Gemoets H, Abraham I, Vlaminck H, Evers G, et al. Incidence, determinants, and consequences of subclinical noncompliance with immunosuppressive therapy in renal transplant recipients. Transplantation 1995 Feb 15;59(3):340-347. [Medline: 7871562]

8. Nevins TE, Kruse L, Skeans MA, Thomas W. The natural history of azathioprine compliance after renal transplantation. Kidney Int 2001 Oct;60(4):1565-1570 [FREE Full text] [doi: 10.1046/j.1523-1755.2001.00961.x] [Medline: 11576374]
9. Vlaminck H, Maes B, Evers G, Verbeke G, Lerut E, Van Damme B, et al. Prospective study on late consequences of subclinical non-compliance with immunosuppressive therapy in renal transplant patients. Am J Transplant 2004 Sep;4(9):1509-1513. [doi: 10.1111/j.1600-6143.2004.00537.x] [Medline: 15307839]
10. Desmyttere A, Dobbels F, Cleemput I, De Geest S. Noncompliance with immunosuppressive regimen in organ transplantation: is it worth worrying about? Acta Gastroenterol Belg 2005 Jul;68(3):347-352. [Medline: 16268422]
11. Butkus DE, Meydrech EF, Raju SS. Racial differences in the survival of cadaveric renal allografts. Overriding effects of HLA matching and socioeconomic factors. N Engl J Med 1992 Sep 17;327(12):840-845. [doi: 10.1056/NEJM199209173271203] [Medline: 1508243]
12. Fan PY, Ashby VB, Fuller DS, Boulware LE, Kao A, Norman SP, et al. Access and outcomes among minority transplant patients, 1999-2008, with a focus on determinants of kidney graft survival. Am J Transplant 2010 Apr;10(4 Pt 2):1090-1107. [doi: 10.1111/j.1600-6143.2009.03009.x] [Medline: 20420655]
13. Chakkera HA, O'Hare AM, Johansen KL, Hynes D, Stroupe K, Colin PM, et al. Influence of race on kidney transplant outcomes within and outside the Department of Veterans Affairs. J Am Soc Nephrol 2005 Jan;16(1):269-277 [FREE Full text] [doi: 10.1681/ASN.2004040333] [Medline: 15563568]
14. Foster CE, Philosophe B, Schweitzer EJ, Colonna JO, Farney AC, Jarrell B, et al. A decade of experience with renal transplantation in African-Americans. Ann Surg 2002 Dec;236(6):794-804; discussion 804. [doi: 10.1097/01.SLA.0000036882.91371.EA] [Medline: 12454518]
15. Eckhoff DE, Young CJ, Gaston RS, Fineman SW, Deierhoi MH, Foushee MT, et al. Racial disparities in renal allograft survival: a public health issue? J Am Coll Surg 2007 May;204(5):894-902; discussion 902. [doi: 10.1016/j.jamcollsurg.2007.01.024] [Medline: 17481506]
16. Opelz G, Mickey MR, Terasaki PI. Influence of race on kidney transplant survival. Transplant Proc 1977 Mar;9(1):137-142. [Medline: 325754]
17. Rostand SG. US minority groups and end-stage renal disease: a disproportionate share. Am J Kidney Dis 1992 May;19(5):411-413. [Medline: 1585926]
18. Weng FL, Israni AK, Joffe MM, Hoy T, Gaughan CA, Newman M, et al. Race and electronically measured adherence to immunosuppressive medications after deceased donor renal transplantation. J Am Soc Nephrol 2005 Jun;16(6):1839-1848 [FREE Full text] [doi: 10.1681/ASN.2004121059] [Medline: 15800121]
19. Kerman RH, Kimball PM, Van Buren CT, Lewis RM, Kahan BD. Stronger immune responsiveness of blacks vs whites may account for renal allograft survival differences. Transplant Proc 1991 Feb;23(1 Pt 1):380-382. [Medline: 1990558]
20. Leffell MS, Steinberg AG, Bias WB, Machan CH, Zachary AA. The distribution of HLA antigens and phenotypes among donors and patients in the UNOS registry. Transplantation 1994 Nov 27;58(10):1119-1130. [Medline: 7974720]
21. Rostand SG, Brown G, Kirk KA, Rutsky EA, Dustan HP. Renal insufficiency in treated essential hypertension. N Engl J Med 1989 Mar 16;320(11):684-688. [doi: 10.1056/NEJM198903163201102] [Medline: 2922014]
22. Krishna S, Boren SA. Diabetes self-management care via cell phone: a systematic review. J Diabetes Sci Technol 2008 May;2(3):509-517. [Medline: 19885219]
23. Logan AG, McIsaac WJ, Tisler A, Irvine MJ, Saunders A, Dunai A, et al. Mobile phone-based remote patient monitoring system for management of hypertension in diabetic patients. Am J Hypertens 2007 Sep;20(9):942-948. [Medline: 17765133]
24. Scherr D, Kastner P, Kollmann A, Hallas A, Auer J, Krappinger H, et al. Effect of home-based telemonitoring using mobile phone technology on the outcome of heart failure patients after an episode of acute decompensation: randomized controlled trial. J Med Internet Res 2009 Aug 17;11(3):e34. [Medline: 19687005]
25. Krishna S, Boren SA, Balas EA. Healthcare via cell phones: a systematic review. Telemed J E Health 2009 Apr;15(3):231-240. [Medline: 19382860]
26. Pare G, Sicotte C, St-Jules D, Gauthier R. Cost-minimization analysis of a telehomecare program for patients with chronic obstructive pulmonary disease. Telemed J E Health 2006 Apr;12:114-121. [Medline: 16620165]
27. Quinn CC, Clough SS, Minor JM, Lender D, Okafor MC, Gruber-Baldini A. WellDoc mobile diabetes management randomized controlled trial: change in clinical and behavioral outcomes and patient and physician satisfaction. Diabetes technology & therapeutics 2008 Jun;10(3):160-168. [Medline: 18473689]
28. Kauer SD, Reif SC, Crooke AH, Khor A, Hearps SJ, Jorm AF, et al. Self-monitoring using mobile phones I the early stages of adolescent depression: randomized controlled trial. J Med Internet Res 2012 Jun;14(3):e67. [Medline: 22732135]
29. Vuong AM, Huber J, Bolin JN, Ory MG, Moudini DM, Helduser J, et al. Factors affecting acceptibility and usability of technological approaches to diabetes self-management: a case study. Diabetes Technol Ther 2012 Dec 14. [Medline: 23013155]
30. 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 Feb 16;14(1):e31. [Medline: 22356799]
31. Boland P. The emerging role of cell phone technology in ambulatory care. J Ambul Care Manage 2007 Apr;30(2):126-133. [Medline: 17495681]

32. Seto E, Leonard KJ, Cafazzo JA, Barnsley J, Masino C, Ross HJ. Perceptions and experiences of heart failure patients and clinicians on the use of mobile phone-based telemonitoring. J Med Internet Res 2012 Feb 10;14(1):e25. [Medline: 22328237]
33. Proudfoot J, Parker G, Pavlovic DH, Manicavasagar V, Adler E, Whitton A. Community attitudes to the appropriation of mobile phones for monitoring and manageing depression, anxiety, and stress. J Med Internet Res 2010 Dec 19;12(5):e64. [Medline: 21169174]
34. Bostock Y, Hanley J, McGown D, Pinnock H, Padfield P, McKinstry B. The acceptability to patients and professionals of remote blood pressure monitoring using mobile phones. Prim Health Care Res Dev 2009 Oct;10(4):299-308. [doi: 10.1017/S1463423609990107]
35. Cleland J, Caldow J, Ryan D. A qualitative study of the attitudes of patients and staff to the use of mobile phone technology for recording and gathering asthma data. J Telemed Telecare 2007;13(2):85-89. [doi: 10.1258/135763307780096230] [Medline: 17359572]
36. Seto E, Leonard KJ, Masino C, Cafazzo JA, Barnsley J, Ross HJ. Attitudes of heart failure patients and health care providers towards mobile phone-based remote monitoring. J Med Internet Res 2010 Nov 29;12(4):e55. [Medline: 21115435]
37. Axelrod DA, Dzebisashvili N, Schnitzler MA, Salvalaggio PR, Segev DL, Gentry SE, et al. The interplay of socioeconomic status, distance to center, and interdonor service area travel on kidney transplant access and outcomes. Clin J Am Soc Nephrol 2010 Dec;5(12):2276-2288. [Medline: 20798250]
38. Axelrod DA, Guidinger MK, Finlayson S, Schaubel DE, Goodman DC, Chobanian M, et al. Rates of solid-organ wait-listing, transplantation, and survival among residents of rural and urban areas. JAMA 2008 Jan 9;299(2):202-207. [Medline: 18182602]
39. Bharmal M, Payne K, Atkinson MJ, Desrosiers MP, Morisky DE, Gemmen E. Validation of an abbreviated Treatment Satisfaction Questionnaire for Medication (TSQM-9) among patients on antihypertensive medications. Health Qual Life Outcomes 2009 Apr 27;7:36. [Medline: 19397800]
40. Morisky DE, Ang A, Krousel-Wood M, Ward HJ. Predictive validity of a medication adherence measure in an outpatient setting. J Clin Hypertens (Greenwich) 2008 May;10:348-354. [Medline: 18453793]
41. Cohen S, Kamarck TMermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [Medline: 6668417]
42. Pinnock H, Slack R, Pagliari C, Price D, Sheikh A. Professional and patient attitudes to using mobile phone technology to monitor asthma: questionnaire survey. Prim Care Respir J 2006 Aug;15(4):237-245. [Medline: 16843066]
43. Spacapan S, Oskamp S. Perceived stress in a probability sample of the United States. In: The social psychology of health. Newbury Park, Calif: Sage Publications; 1988:31-67.
44. Pare G, Moqadem K, Pineau G, St-Hilaire C. Clinical effects of home telemonitoring in the context of diabetes, asthma, heart failure and hypertension: a systematic review. J Med Internet Res 2010 Jun 16;12(2):e21. [Medline: 20554500]
45. Mistry H. Systematic review of studies of the cost-effectiveness of telemedicine and telecare. Changes in the economic evidence over twenty years. J Telemed Telecare 2012 Jan;18(1):1-6. [Medline: 22101609]
46. Smith A. Americans and their cell phones. Pew Internet & American Life Project 2011 Aug 15 [FREE Full text]
47. Nielsen Wire. 2012 Mar 29. Smartphones account for half of all mobile phones, dominate new phone purchases in the US URL: http://blog.nielsen.com/nielsenwire/online_mobile/smartphones-account-for-half-of-all-mobile-phones-dominate-new-phone-purchases-in-the-us [accessed 2012-07-23] [WebCite Cache ID 69NNChbUD]

Abbreviations

HTN: hypertension
mHealth: mobile phone based health technology

Edited by G Eysenbach; submitted 26.07.12; peer-reviewed by A Rangarajan; comments to author 13.09.12; revised version received 05.10.12; accepted 05.10.12; published 08.01.13.

Please cite as:
McGillicuddy JW, Weiland AK, Frenzel RM, Mueller M, Brunner-Jackson BM, Taber DJ, Baliga PK, Treiber FA
Patient Attitudes Toward Mobile Phone-Based Health Monitoring: Questionnaire Study Among Kidney Transplant Recipients
J Med Internet Res 2013;15(1):e6
URL: http://www.jmir.org/2013/1/e6/
doi:10.2196/jmir.2284
PMID:23305649

Original Paper

Mobile Health (mHealth) Approaches and Lessons for Increased Performance and Retention of Community Health Workers in Low- and Middle-Income Countries: A Review

Karin Källander[1*], MSc, PhD; James K Tibenderana[1*], MB ChB, PhD; Onome J Akpogheneta[2*], BSC, PhD; Daniel L Strachan[3*], BA, MSc; Zelee Hill[3*], MA, MSc, PhD; Augustinus H A ten Asbroek[4*], RN, MSc, PhD; Lesong Conteh[4*], BSC, MSc, PhD; Betty R Kirkwood[4*], MA, MSc, DIC, FFPH, FMedSci; Sylvia R Meek[2*], MA, MSc, PhD

[1]Malaria Consortium Africa, Kampala, Uganda

[2]Malaria Consortium, London, United Kingdom

[3]Institute of Global Health, University College London, London, United Kingdom

[4]Faculty of Epidemiology and Population Health, Department of Population Health, London School of Hygiene and Tropical Medicine, London, United Kingdom

[*]all authors contributed equally

Corresponding Author:
Karin Källander, MSc, PhD
Malaria Consortium Africa
P.O Box 8045
Kampala
Uganda
Phone: 256 772744126
Fax: 256 312 300425
Email: k.kallander@malariaconsortium.org

Abstract

Background: Mobile health (mHealth) describes the use of portable electronic devices with software applications to provide health services and manage patient information. With approximately 5 billion mobile phone users globally, opportunities for mobile technologies to play a formal role in health services, particularly in low- and middle-income countries, are increasingly being recognized. mHealth can also support the performance of health care workers by the dissemination of clinical updates, learning materials, and reminders, particularly in underserved rural locations in low- and middle-income countries where community health workers deliver integrated community case management to children sick with diarrhea, pneumonia, and malaria.

Objective: Our aim was to conduct a thematic review of how mHealth projects have approached the intersection of cellular technology and public health in low- and middle-income countries and identify the promising practices and experiences learned, as well as novel and innovative approaches of how mHealth can support community health workers.

Methods: In this review, 6 themes of mHealth initiatives were examined using information from peer-reviewed journals, websites, and key reports. Primary mHealth technologies reviewed included mobile phones, personal digital assistants (PDAs) and smartphones, patient monitoring devices, and mobile telemedicine devices. We examined how these tools could be used for education and awareness, data access, and for strengthening health information systems. We also considered how mHealth may support patient monitoring, clinical decision making, and tracking of drugs and supplies. Lessons from mHealth trials and studies were summarized, focusing on low- and middle-income countries and community health workers.

Results: The review revealed that there are very few formal outcome evaluations of mHealth in low-income countries. Although there is vast documentation of project process evaluations, there are few studies demonstrating an impact on clinical outcomes. There is also a lack of mHealth applications and services operating at scale in low- and middle-income countries. The most commonly documented use of mHealth was 1-way text-message and phone reminders to encourage follow-up appointments, healthy behaviors, and data gathering. Innovative mHealth applications for community health workers include the use of mobile phones as job aides, clinical decision support tools, and for data submission and instant feedback on performance.

Conclusions: With partnerships forming between governments, technologists, non-governmental organizations, academia, and industry, there is great potential to improve health services delivery by using mHealth in low- and middle-income countries. As with many other health improvement projects, a key challenge is moving mHealth approaches from pilot projects to national scalable programs while properly engaging health workers and communities in the process. By harnessing the increasing presence of mobile phones among diverse populations, there is promising evidence to suggest that mHealth can be used to deliver increased and enhanced health care services to individuals and communities, while helping to strengthen health systems.

(J Med Internet Res 2013;15(1):e17) doi:10.2196/jmir.2130

KEYWORDS

mHealth; community health worker; Africa

Introduction

Community health workers were a cornerstone of primary health care as envisaged by the Declaration of Alma-Ata in 1978, yet the enthusiasm for community health workers started to diminish by the early 1990s, partly because of the challenges of scaling-up programs in a sustainable fashion while maintaining their effectiveness [1]. However, due to slow progress toward the United Nations' Millennium Development Goals, community-based programs delivering care to sick children have yet again become priorities to curb child mortality in high-mortality countries [2]. Integrated community case management (ICCM) of malaria, pneumonia, and diarrhea delivered by community health workers is such a strategy, which is now being implemented at scale in several African countries, including Uganda and Mozambique. The purpose of ICCM is to improve access to effective treatment for sick children among hard-to-reach populations with the ultimate goal of reducing under-5 mortality.

The Innovations at Scale for Community Access and Lasting Effects (inSCALE) project, a collaboration between Malaria Consortium, London School of Hygiene and Tropical Medicine, and University College of London, aims to better understand community health workers motivation and supervision, and to find feasible and acceptable solutions to community health workers' retention and performance, both of which are vital for successful implementation of ICCM at scale. One such solution could be to use mobile phones as a tool to increase the status of community health workers in the community and allow frequent feedback and support to community health workers based on data submitted, potentially resulting in improved quality of care delivered.

With almost 5 billion mobile phone users in the world, health care providers and researchers are realizing the potential of using mobile technologies, such as mobile phones, portable computers, and personal digital assistants (PDAs), for health services. Mobile health (mHealth), as defined by the World Health Organization (WHO), is an area of electronic health (eHealth) that provides health services and information via mobile technologies such as mobile phones and PDAs. mHealth can also support the performance of health care workers by the dissemination of clinical updates, learning materials, and reminders [3], particularly in underserved rural locations in low- and middle-income countries where community health workers deliver ICCM to children sick with diarrhea, pneumonia, and malaria [3,4].

The aim of this review is to provide a thematic overview of mHealth project approaches to the intersection of mobile technology and public health and the application of these approaches in programs specifically focusing on community health workers. The potential challenges and opportunities for integration of such mHealth applications in existing national systems are discussed.

Methods

A 2-stage process was applied in this review. In the first stage, a broad search was done to generate a list of domains in which mHealth has been applied in low- and middle-income countries. Non-peer-reviewed sources of information, including Web-based mHealth portals, mHealth review documents [4-7], and reports that specifically map out mHealth initiatives in Uganda and Mozambique [8,9] were used (Table 1). The search in this stage did not aim to generate a comprehensive list of all mHealth projects conducted in low- and middle-income countries; the purpose was to generate different mHealth domains and give examples of projects that have addressed the domains identified.

Table 1. Sources used to identify mHealth projects.

Source	Name
mHealth portals	The Communication Initiative Network [10]
	KIT Royal Tropical Institute [11]
Review documents	Mechael et al [5]
	Vital Wave Consulting [6]
	Vital Wave Consulting [7]
Country reports	Macueve [8]
	Mwagale and Kakaire [9]

In the second review stage, a narrower and more systematic search was done to identify projects that applied mHealth for community health workers in low- and middle-income countries. In addition to the sources provided in Table 1, peer-reviewed papers were identified from searches on PubMed and Google Scholar by using the following search terms: mHealth, mobile health, developing, and low income. Projects conducted in countries with low, lower-middle, or upper-middle incomes (as per World Bank definition) qualified for inclusion if they mentioned that at least 1 of their user groups included community health workers. Although the main reviews for both stages 1 and 2 were conducted by the first author (KK) between June and September 2010, the review in stage 2 was updated in April 2012 by using the same search terms as used previously. Project descriptions published after these dates are not included in this review.

The information from stage 1 was structured by the first author (KK) who classified the projects into predefined themes. Various numbers of themes have been used previously within the mHealth ecosystem, ranging from 4 different categories in a review by Blynn [12], 5 in a paper by Mechael et al [5], and 6 in a report by Vital Wave Consulting [6]. To give a more diversified description of the projects reviewed, we adopted the 6-theme classification used by Vital Wave Consulting (Table 2).

Table 2. The six themes in mHealth.

Theme[a]	Description
Education and awareness	Primarily 1-way communication programs to mobile subscribers via SMS/text messaging in support of public health, behavior-change campaigns.
Data access	Applications designed to use mobile phones, PDAs, or laptops to enter and access patient data. Some projects may also be used by patients to access their own records.
Monitoring and compliance	One- or 2-way communications to the patient to monitor health conditions, maintain caregiver appointments, or ensure strict medication regimen adherence. Some applications may also include inpatient and outpatient monitoring sensors for monitoring of multiple conditions (such as diabetes, vital signs, or cardiac).
Disease and emergency tracking	Applications using mobile devices to send and receive data of disease incidence, outbreaks, geographic spread of public health emergencies, often in association with Global Positioning System (GPS) systems and backend applications for visualization.
Health information systems	Applications developed for "back office" or central health care information technology systems allowing for access by and integration with mHealth application. Such applications often tie-in to regional, national, or global systems.
Diagnosis and consultation	Applications developed to provide support for diagnostic and treatment activities of remote caregivers through Internet access to medical information databases or to medical staff.

[a] Adopted from Vital Wave Consulting [6].

Key mHealth technologies reviewed included mobile phones, PDAs and smartphones, patient monitoring devices, and mobile telemedicine devices. The ways these tools can be used for education and awareness, data access, and for strengthening health information systems were explored. We also considered how mHealth may support patient monitoring, clinical decision making, and tracking of drugs, supplies, and emergencies.

For information gathered in stage 2 of the review process, the results were analyzed, classified into a theme, and described in more detail by using a predesigned data collection table (available on request from first author). The results were presented to the coauthors in an 8-person research team meeting and the classification was discussed until consensus was achieved.

Results

Common mHealth Solutions

The main capabilities provided by mHealth applications are voice, text, and data access with information going 1 way, 2 ways, or multiple ways (Table 3). These applications included phone calls (personal or automated robocalls with or without toll-free numbers), text messages (including personal text reminders or mass texting for community mobilization), data transfer for health record tracking or clinical decision support, and mobile telemedicine devices for patient monitoring or diagnosis.

Table 3. Example of mHealth applications utilizing 1-way, 2-way, or multiway communication.

Communication type	Example mHealth application
1-way	SMS/rapid SMS
	Sending data
	Push messages
	“Please call me” message
2-way	Sending data and receiving feedback
	Quizzes/games
	Hotlines/textlines
	Remote consultation/training
Multiway	Frontline SMS
	Facebook
	Twitter

One-Way, Two-Way, or Multiway Applications

Communication between a sender and receiver can occur in more than 1 direction and within varying group sizes. One-way communication is similar to mass media that distributes information in 1 direction. mHealth innovations have typically been designed as 1-way communications in which projects use “push” technology to deliver information to subscribers’ phones by using messages tailored to personal needs. Most commonly identified push designs include bulk short message service (SMS) or robocalls to large audiences.

Two-way communication is interactive and more similar to interpersonal communication. For users, interactivity may require greater effort and generate greater interest. Interactive quizzes, information menus, data collection and tailored responses, hotlines, and interactive voice responses are examples of 2-way communication [5]. Although most 2-way communication does not occur in real time, some applications, such as closed user groups or voice over IP (VoIP) for remote health consultations and health worker training, do use real-time communication [6].

Multiway communication can vary the number of senders and receivers, including 1-to-many, many-to-1, and many-to-many communication. Many-to-many communications include social media applications, such as Facebook or Twitter, that can be accessed from most Internet-enabled mobile phones. Most mHealth projects used a combination of 1-way and 2-way communication methods pertinent to several themed categories in Table 2, whereas only a few projects could be identified that used social media.

Education and Awareness

The Cellphones4HIV project in South Africa described by de Tolly and Alexander [13] sends out messages on antiretroviral treatment adherence using Unstructured Supplementary Service Data (USSD) (ie, the system used to load airtime), Mxit (a Java application that allows general packet radio service [GPRS] or 3G-based instant messaging) and voicemail messages pushed into the user’s voicemail inbox with notification by SMS. Push designs were found to have differing capabilities, limitations, and requirements, but may be combined, adapted, or further expanded as technology evolves.

Projects for remote health information dissemination, like Project Masiluleke [14] and Text-to-Change (TTC) [15], have reached large audiences with information on HIV prevention and treatment using ”please call me” (PCM) messages and bulk SMS. PCM messages have been widely used in mHealth projects in Africa because they are free for senders and can be sent from phones that have no credit. Project Masiluleke in South Africa sent 1 million PCM messages per day for 1 year, offering contact information for local HIV and tuberculosis call centers [14].Within 5 months, calls to South Africa’s National AIDS helpline quadrupled [16]. In Uganda, TCC used a bulk SMS platform to create dialog and increase awareness of HIV in order to reduce related stigma and discrimination, and motivate people to seek HIV testing and treatment [15]. TTC also sent out quizzes and information about HIV prevention and testing, awarding those who pass the quiz with airtime. Of 15,000 subscribers contacted by TTC, 2500 responded to each question.

In FHI360-SATELLIFE’s Uganda Health Information Network (UHIN) project, continuing medical education targeted to health workers was broadcast 3 times per week via PDAs regarding diagnosis, treatment, and prevention of major health problems [17]. In addition, they received daily news from mainstream media. Other projects used SMS for behavior-change communication. The Text2Teach project gave Philippine teachers a mobile phone texting platform to receive videos via satellite over school-based televisions and mobile technology involving parents [18]. Behavior-change communication can be used in various applications, from family planning and teenage pregnancy to disease awareness and prevention to advice on agricultural and farming techniques.

Social networks, such as Facebook, Twitter, or Hi5, are used by hundreds of millions of people to communicate about a huge range of topics, including health. The WHO used Twitter during the influenza A (H1N1) pandemic and, at time of writing, had more than 11,700 followers from all over the world [19]. In Mozambique, the nonprofit organization DKT International launched a social franchising program, branded as Intimo, that uses social media to increase access to its clinics. Its Facebook page reaches over 6600 Mozambicans (85% between the ages of 18 and 34 years) with information on family planning and reproductive health [20]. In Indonesia, the Fiesta condom brand has used Facebook, Twitter, and YouTube to talk about safe sex and condom use [21].

Community Health Worker Program Innovations for Education and Awareness

Through SMS with community members and community health workers, mHealth has opportunities to communicate health messages directly and simultaneously [22]. The SMS campaigns for health education, promotion, and awareness typically used SMS to disseminate information and prevention messaging or direct patients to services. Mobile phones also present opportunities for community health workers to communicate directly with one another and provide peer support [4]. To provide additional support to community health workers during home visits, the Tanzanian Mobile Video for Community Health Workers project used the CommCare tool to provide health education videos played on mobile phones [23].

Stakeholders suggested expanding 1-way to 2-way communications, including introducing a referral alert process in which community health workers call health facilities before the patients' arrival [4]. Establishing call-in services for each health facility could also allow community health workers to receive updated information on drug stocks, attendance records, and other relevant information. In addition, appointment confirmation texts for referred patients with time, date, and appointment location could be effective, as well as SMS alerts to community health workers about appointments attended by referred patients. Texts or SMS could also be used by health facility workers and community health workers to keep each other informed of recent developments and upcoming events, including SMS to community health workers on their birthday for motivation [4].

The concern that national privacy laws can hinder projects from accessing the target beneficiaries' personal phone numbers was raised. One stakeholder mentioned a project in which a collection of mobile phone numbers for health workers to send push messages had to be stopped after concerns were raised about the assumption that all health providers had given their permission to allow projects to reach them on their telephones (J Tibenderana, personal communication, September 2010).

Data Access

Innovations in mHealth can conceivably change how data are used in health programs, leading to faster, decentralized decision making and reallocation of resources due to faster data analysis [22]. Handheld computers, PDAs, or laptops for data collection and reporting can use 1- or 2-way communication systems. RapidSMS has established a 2-way flow of communication that empowers stakeholders with a dynamic tool for fast, efficient, and accurate data collection, analysis, and communication [24]. In addition, SMS-based data for health care workers can identify, diagnose, and track patients by using streamlined technology that is automatically updated in a central system.

Twelve Ugandan projects used mobile technology for data collection and reporting [9]. Most were designed as 1-way communication systems to improve data collection or management in surveys, routine care, and vaccine trials.

Community Health Worker Program Innovations for Data Access

Although there is little evidence of the effectiveness of community health workers collecting and self-reporting data from patient records, mobile phones have been suggested as a useful tool for rural health workers' reporting of data as it is suggested it improves accuracy, reduces time and cost, and improves data quality [19]. A cost-effectiveness study showed that using PDAs for data collection delivered 24% savings per unit of spending over traditional manual data collection and transmission approaches [25]. However, use of PDAs in a Rwandan ICCM program exacerbated, rather than lessened, volunteer workload [4]; mobile phone-assisted data collection became onerous and was felt to have distanced community health workers from the human side of their role, turning them into "data collection robots."

Blaschke et al [26] and the Millennium Villages Project [27] describe the use of ChildCount+ that uses mobile technologies for improving data use and reporting among community health workers in several African countries, including Malawi and Uganda. This platform, developed by the Millennium Villages Project, aimed to improve maternal and child survival by supporting delivery of community-based management of acute malnutrition, malaria, and diarrhea. Three months after initiation, 95% of 9561 children under 5 years in the Malawian cluster had been registered using mobile technology, and only approximately 10% of incoming messages to the system were rejected due to improper formatting [26]. The RapidSMS platform used led to significant reduction in data transmission delay compared to Malawi's current paper-based system.

Monitoring and Compliance

Text messaging via mobile phones has garnered increasing attention as a means of reminding patients of appointments in the United Kingdom, United States, Norway, and Sweden. This resulted in a lowering of nonattendance to scheduled appointments, yielding significant savings in health costs for facilities and practitioners [28]. In this case, the benefit is cost-related rather than health outcome-related.

In addition, SMS has also been used as a way of monitoring patients' medication compliance. However, literature on treatment compliance has focused primarily on management of chronic diseases, such as diabetes, smoking cessation, and breast cancer, in high-income countries and few examples exist from low- and middle-income countries [5]. A South African trial showed tuberculosis patients with increased compliance rates, and a Thai study showed that 90% of tuberculosis patients receiving daily SMS medication reminders adhered to treatment [7,12]. A Kenyan efficacy study provided 428 HIV patients with mobile phones and randomized patients to receive daily, weekly, or no SMS reminders. Treatment adherence was improved for patients receiving weekly, but not daily, SMS and treatment interruptions were less likely [29]. Adding words of encouragement to an SMS did not prove more effective and confidentiality was a concern.

To improve medicine compliance and adherence to antiretroviral drugs in Uganda, a medical container called Wisepill was used to transmit a cellular signal whenever opened, send weekly SMS at preset times, and provide interactive voice response [30]. A similar project, SIMpill, monitored adherence to tuberculosis drugs in South Africa [31]. Few randomized controlled trials studying treatment compliance were found, and statistically significant results were limited by sample size; mixed results have been found in other studies [32]. A strong focus on feasibility and usability was evident, with little connection to health outcomes [5].

Other mHealth applications can be used to improve compliance to guidelines by health workers. A proof-of-concept randomized controlled Kenyan trial on adherence to malaria treatment guidelines used 10 carefully designed SMSs with drug delivery instructions and an unrelated motivational message to aid rural health facility workers [33]. Both immediate and 6-month analyses showed improved malaria case management. The trial is undergoing cost-effectiveness analysis and qualitative analysis to examine possible added burdens on health workers.

Community Health Worker Program Innovations for Monitoring Compliance

A randomized controlled trial delivered SMS to community-based peer health workers in rural Uganda supporting antiretroviral treatment for HIV patients [34]. No virological differences in patient outcome over 26 months were observed, but limited qualitative data showed improvements in patient care, logistics, and broad support from health workers and patients. Improvements in peer health worker morale and confidence were reported; peer health worker-patient relationships improved, shifting burdens from staff-patient relationships. As compared to voice calls, reservations about the lack of immediate response via SMS were noted, privacy concerns were raised, and phone maintenance and charging were also problematic.

Disease and Emergency Tracking

Several countries have used mHealth innovations for not only disease tracking, but also for supply tracking. The Foundation for Innovative New Diagnostics (FIND) deployed RapidSMS in 2 districts in Uganda and worked with health centers to submit and map weekly epidemiological records, malaria case management, and malaria medicine stock reports [35]. The platform EpiSurveyor has also been widely used for emergency response and tracking supplies. It allows users to download, fill, and send forms to central databases for real-time analysis [36].

Mobile phones and Web-based technologies have also been used for early warning of disease outbreaks. The Acute Encephalitis Syndrome Surveillance Information System (AESSIMS) project in India aimed to improve immunization services for Japanese B encephalitis, diphtheria, hepatitis B, measles, pertussis, tetanus, and polio by tracking diseases in real time [37].

Reports have described mobile technology use during natural disasters, including the earthquakes in China in 2008 and Haiti in 2010 [38,39]. Mobile phones were primarily used for tracking population movements, infectious disease reporting, and coordinating search and rescue missions. Studies investigating mobile phone use for telemedicine during emergencies found them effective for relatively fast and accurate in-transit patient treatment, sending images for diagnosis, and using video capabilities.

Community Health Worker Program Innovations for Disease and Emergency Warning Systems

As part of Cambodia's malaria elimination strategy, the National Center for Parasitology, Entomology and Malaria Control (CNM), with technical support from Malaria Consortium and WHO, village malaria workers are trained to send SMSs to report malaria cases in real time [40]. These SMS messages also support the paper reporting that feeds into the health information system from the health centers. The project had low start-up costs, estimated at US $100 for each village malaria worker, which includes a mobile phone, subscriber identity module (SIM) card, solar charger, and training. Because of the effective cooperation with the private sector, all SMS messaging is free resulting in essentially zero maintenance costs [41].

In areas where outbreaks of disease occur, community health workers could use mHealth to track medicine stocks (eg, FIND) and report observed cases with daily case statistics delivered using FrontlineSMS [40]. Community health workers can also minimize the impact of outbreaks by disseminating educational information about disease prevention and handling. In the Healthy Child Uganda project, community health workers used mobile phones to send emergency alerts and requisition supplies to support ICCM activities in treating pneumonia, diarrhea, and malaria [42].

Health Information Systems

Health administration systems are used for epidemiological research, tracking of indicators for monitoring and evaluation, and financial and cost reporting for supply management [6]. Mozambique used PDAs to support collection of data from health records [43]. The stand-alone system, known as "módulo básico," has now been implemented in all provinces and districts in the country [44].

Several African countries, including Mozambique and Uganda, have tested 2-way access to district health information by using mobile phone networks and low-cost PDAs for data dissemination, collection and reporting, and email exchange [17,45]. The Mozambique Health Information Network (MHIN) set up data transfer via PDAs using wireless access points and a server located at the Ministry of Health in Mozambique. District health offices received data from health centers and used the network to monitor drug stocks and guide orders. Up to 50% improvement in data quality was observed. The MHIN services are expanding to additional districts and cost-benefit analyses comparing MHIN- and paper-based approaches are planned [45].

The same team who worked on MHIN also set up UHIN in Uganda [17]. Health workers used PDAs to collect and upload data and emails via infrared, Bluetooth, or Wi-Fi at rural health facilities. The access point sent data and messages via mobile networks to the server, which routed them to the correct recipients and sent return messages with data and health information.

The public-private SMS for Life project in Tanzania used mobile phones and electronic mapping technology to generate and deliver weekly information to health centers on malaria medicines [46]. The project proved successful, and medicine stock-out rates were significantly reduced within 21 days.

Sustainability of countrywide mHealth programs relies on incorporation with the national health care program of the country, yet few African countries have developed national eHealth or mHealth policies, strategies, or guidelines [5]. Much of this is because of the limited knowledge of what works, how it works, and how much it costs. An exception is Ethiopia, where a national policy for eHealth is about to be launched [47].

Community Health Worker Program Innovations for Health Information Systems

Few studies have examined health information and administration systems that include community health workers. The ICT4MPOWER project is a 3-year proof-of-concept project in Uganda aiming to increase health system effectiveness and empower community health workers in rural areas by aiding referrals and patient follow-up, while ensuring transfer of skills and knowledge to health workers [48]. The Tanzanian CommCare project provided a community health mobile platform, enabling community health workers to provide more efficient care and to receive better supervision [49]. Such projects indicate the great potential to link community health workers with health administration systems by using mobile technologies that would add value to government health policy, providing integrated health data and a dynamic picture of national health care provision.

Diagnosis and Consultation

Use of electronic technologies to provide support for diagnosis, consultation, and treatment activities conducted by remote caregivers is increasingly common. Mobile phones can be used as respiratory or pulse rate counters, gestational age date calculators, drug dose calculators, drip rate calculators, and drug reminder alarms when installed in mobile phones and linked to a sensor [50]. Another example of a diagnostic tool is CellScope, which uses a modified mobile phone for blood, urine, or other sample loading for malaria, HIV, and tuberculosis diagnosis [51]. None of these applications requires any transfer of data; hence, running costs are close to zero.

A pilot study of Electronic Integrated Management of Childhood Illnesses (eIMCI) in rural Tanzania, tested whether PDAs could improve diagnosis of children using IMCI protocols. The project was found to be feasible and acceptable to health workers in providing mobile decision support [52]. In addition, 6 Ugandan projects used mobile phones to send medical test results through SMS or email to patients and health workers; others used wireless devices to provide clinical training and patient care support services [9].

Community Health Worker Program Innovations for Diagnosis and Consultation

RapidSMS can be used in various ways, including supporting community health worker-patient interactions [24]. Mobile phones used as job aides could allow community health workers, via SMS or data transfer, to send patient information and receive instructions on how to proceed [27]. This could demonstrate program effectiveness to community health workers, potentially motivating continued work and better service [4]. In Colombia, the CellPhone GuideView system broke down complex diagnostic and treatment procedures into simple steps for community health workers using an authoring tool in which text, pictures, audio, and video were embedded to aid comprehension and ease of use [50,53]. Community health workers were then able to transmit images, data, and audio to remote experts for further advice.

Discussion

The review revealed that there are very few formal outcome evaluations of mHealth in low-income countries. Although there is vast documentation of project process and uptake, most were evaluations of small-scale pilot studies that were not designed to demonstrate an impact on behavior change or health. There is also a lack of mHealth applications and services operating at scale in low- and middle-income countries. The most commonly documented use of mHealth was 1-way text-message and phone reminders to encourage follow-up appointments, healthy behaviors, and data gathering. Two-way communication applications focused primarily on data transmission with automated feedback response, and few projects were implementing real-time communication. Although some claim that social media can be an effective tool for engaging patients online [54], others argue that health institutions need to develop clear policies about the use of social media in patient care environments to ensure patient safety [55]. However, the majority of multiway and social media projects identified in this review were patient/user driven, such as Facebook or Twitter, with little or no involvement of treating physicians or nurses.

A limited number of mHealth projects were found which specifically targeted community health workers. Of the few projects identified, most used a combination of simple mobile phone applications for data submission, job aids to improve diagnostics, and for sending and receiving SMS messages and reminders. None of these projects had evaluated the impact of these tools on community health workers quality of care provided. Most projects used applications that communicated by using 1-way or 2-way SMS, whereas GPRS-enabled applications were rare. Although several projects tested applications that aimed to improve accuracy in community health worker data submission and clinical decision-making skills using electronic job aids [26,27,49], international stakeholders cautioned that these may result in community health workers focusing more on the technology than on the patient [4].

The key considerations for successful use of or expansion of mHealth innovations include collaboration, financing, literacy and cultural, partnerships, and technical considerations (Table 4). As a young field, mHealth is well positioned to benefit from best practices and available technology documented in various project reports. Sustainability and scalability are still the main challenges to the strategic deployment of mHealth applications, partly reflecting the gap between what application developers are doing on the ground and what the governments see as priorities and initiatives they need to step in and support [2]. Establishing true partnerships with users and policy makers throughout the design and implementation processes is critical for success and collaboration with operators could ensure technical support, make scale-up possible, and reduce costs to drive mHealth demand and innovation [4,5]. This is illustrated by examples from Ghana and Cambodia, where physicians registered with the Ghana Medical Association have access to unlimited calls through the mobile service operator, OneTouch [3], and where village malaria workers in Cambodia report malaria cases by using free SMS with Mobitel [40].

Table 4. Key considerations for successful use and expansion of mHealth innovations.

Areas of consideration	Description
Collaboration	Collaborative projects allow simpler widespread implementation
	Collaboration with operators could ensure technical support and make program scaling possible
	Collaboration more likely when all partners display strong affinity to the goal
Costs and sustainable financing	Collaborations can provide resources and support for project costs
	Organizations more likely to commit resources for piloting new initiatives when projects lasted for limited timeframe, when partners maintain control over deliverables, and when funds do not need approval and transference to third party
	Average SMS cost does not exceed US $0.05
	Lost phones and hardware can be mitigated by providing cheaper phones and ensuring equipment bears the program logo
Literacy and cultural specificity	Illiteracy is an important consideration for text-based innovations in low- and middle-income countries
	Accessibility to target users/patients must consider cultural sensitivities
Health worker partnership, engagement, training, and compliance	Partnership with users can enhance design and implementation of projects
	Using iterative cycles with target users when developing a software can stimulate ownership and enhance project engagement
	Data feedback reports should be distributed to users submitting data regularly (perhaps initially once per week)
	Use brief and personal SMS messages, allow opt out, allow language choice (with careful translation), and validate content with target users
Technical considerations	Use user-friendly and project-appropriate equipment
	Maintain 160 character length for SMS
	Take care with abbreviations, slang, and tone in SMS messages
	Ensure enough time for procuring and establishing necessary equipment and phone lines
	First resolve lack of telephone and Internet connectivity among target health care providers
	Use of existing technology, such as “please call me” messages, rather than introduction of new technologies
	Implement clear data usage and storage guidelines with data quality checks and backups made frequently

The national ownership of mHealth applications cannot be overemphasized. Some good examples of country ownership exist, such as state programs in Ghana and Nigeria, which address maternal and neonatal health using mobile phones [56]. The challenge is to have health ministers and officials at the same table as mobile service providers, doctors, technologists, and financiers. Coordination among these stakeholders and agreement of incentive structures and responsibilities for meaningful collaboration is needed to better inform public and private investments and the deployment of commercially viable solutions [5].

The mHealth interventions often used SMS to provide information, motivate individuals, and encourage self-management or promote disease prevention. However, illiteracy is an issue for text-based prevention interventions [5]. Culturally specific provision of health information is important because poorly designed campaigns can have negative unintended effects; good understanding of cultural context and strategies to overcome language and literacy barriers are needed.

As with other mHealth applications, there is a significant gap in evidence on behavioral, social, economic, and health outcomes of using mobile phones and SMS for improving health in low- and middle-income countries, as demonstrated in a systematic review of the literature [57].

Funding in low- and middle-income countries is not adequate to support complex telemedicine in emergency situations. Infrastructural limitations, such as network capacity, also constrain the effectiveness of emergency monitoring and tracking [5]. However, routine data from all active SIM cards could be used in disaster-affected areas for near real-time monitoring of population movements during disease outbreaks [39]. Another significant barrier to implementation of mHealth systems is in relation to health worker resistance to new technology and broader discussion and research about health worker benefits, and incentives for use and compliance is required [5,6,22]. This would include ensuring adequate training also remains a critical component for large-scale implementation [5].

Addressing security and privacy issues in mHealth has also proven challenging. Guidelines on the rights to data, usage, and storage must be outlined and implemented, with sufficient qualitative data to explain potential findings collected alongside close program monitoring. For mHealth success, cooperation between local communities and regional and national health information systems is essential [3,58,59]. It is also unclear from the review whether SMS projects for health workers need to comply with any national privacy laws because collecting health workers' private phone numbers to push messages is assuming that they have all given their permission to have the project reach them on their phones.

Limitations related to the landscape analysis should be considered when interpreting the results. The review focused only on 6 major thematic areas for mHealth and it is possible that some mHealth applications and tools have been excluded. Given the bulk of projects piloting mHealth applications in low- and middle-income countries, the first stage of the review only describes a sample of projects and applications tested under each thematic area. However, the second stage of the review, ie, that of mHealth projects targeting community health workers, was deemed systematic and comprehensive. The sources of the information reviewed were primarily obtained from project websites because few peer-reviewed evaluations were identified, potentially resulting in overreporting of positive results and underreporting of challenges or failures.

With partnerships forming between governments, technologists, non-governmental organizations, academia, and industry, there is great potential to improve health services delivery using mHealth in low- and middle-income countries. As with many other health improvement projects, a key challenge is moving mHealth approaches from pilot projects to national scalable programs while properly engaging health workers and communities in the process. By harnessing the increasing presence of mobile phones among diverse populations, there is promising evidence to suggest that mHealth can be used to deliver increased and enhanced health care services to individuals and communities, while helping to strengthen health systems.

Acknowledgments

We are grateful to the Bill & Melinda Gates Foundation for sponsoring this research under the Innovations at Scale for Community Access and Lasting Effects (inSCALE) project.

Conflicts of Interest

None declared.

References

1. The World Bank. World Development Report 1993: Investing in Health. New York: Oxford University Press; 1993. URL: http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/1993/06/01/000009265_3970716142319/Rendered/PDF/multi0page.pdf [accessed 2012-11-09] [WebCite Cache ID 6C2iQ6eDY]
2. United Nations Development Program. Mobile technologies and empowerment: enhancing human development through participation and innovation. 2012. URL: http://www.undpegov.org/sites/undpegov.org/files/undp_mobile_technology_primer.pdf [accessed 2012-04-04] [WebCite Cache ID 66fbQ4SE6]
3. Mechael PN. The case for mHealth in developing countries. Innovations: Technology, Governance, Globalization 2009;4:103. [doi: 10.1162/itgg.2009.4.1.103]
4. Strachan DL, Källander K, Ten Asbroek AH, Kirkwood B, Meek SR, Benton L, et al. Interventions to improve motivation and retention of community health workers delivering integrated community case management (iCCM): stakeholder perceptions and priorities. Am J Trop Med Hyg 2012 Nov;87(5 Suppl):111-119. [doi: 10.4269/ajtmh.2012.12-0030] [Medline: 23136286]
5. Mechael PN, Batavia H, Kaonga N, Searle S, Kwan A, Goldberger A, et al. Barriers and Gaps Affecting mHealth in Low and Middle Income Countries: Policy White Paper. 2010 May. URL: http://www.mobileactive.org/files/file_uploads/mHealth_Barriers_White_Paper.pdf [accessed 2012-04-04] [WebCite Cache ID 66fZH7CWL]
6. Vital Wave Consulting. mHealth in the Global South Landscape Analysis. 2008. URL: http://www.vitalwaveconsulting.com/pdf/mHealth_in_the_Global_South_Landscape_Analysis.pdf [accessed 2012-04-04] [WebCite Cache ID 66fqSMyQ7]

7. Vital Wave Consulting. mHealth for Development: The Opportunity of Mobile Technology for Healthcare in the Developing World. 2009. URL: http://www.vitalwaveconsulting.com/pdf/mHealth.pdf [accessed 2012-04-04] [WebCite Cache ID 66fqYCKby]
8. Macueve G. eHealth and mHealth in Mozambique: innovations likely to lead to increased retention and performance of Community Based Agents (CBAs). 2010 Sep. URL: http://www.malariaconsortium.org/inscale/downloads/1210-mhealth-mozambique.pdf [accessed 2012-11-09] [WebCite Cache ID 6C2sXmaGH]
9. Mwagale S, Kakaire F. Mobile Health Projects in Uganda. 2010. URL: http://www.malariaconsortium.org/inscale/downloads/1206-mhealth-projects-in-uganda.pdf [accessed 2012-11-09] [WebCite Cache ID 6C2sKj1np]
10. ICT for Development Network. 2010. ICT for Development URL: http://www.comminit.com/ict-4-development/category/sites/global/ict-development [accessed 2012-04-03] [WebCite Cache ID 66eOqqe63]
11. Royal Tropical Institute: mHealth in Low-Resource Settings. 2010. mHealth projects: examples from low- and middle-income countries URL: http://www.mhealthinfo.org/projects_table [accessed 2012-04-03] [WebCite Cache ID 66eNrAn5w]
12. Blynn E. Piloting mHealth: A Research Scan. Cambridge, MA: Knowledge Exchange Management Sciences for Health; 2009. URL: https://wiki.brown.edu/confluence/download/attachments/9994241/mHealth+Final.pdf?version=1&modificationDate=1264456692000 [WebCite Cache ID 66fSKmdIW]
13. de Tolly K, Alexander H. Soul Beat Africa-HIV/AIDS. 2009 Mar 01. Innovative use of cellphone technology for HIV/AIDS behaviour change communications: 3 pilot projects URL: http://www.comminit.com/?q=hiv-aids-africa/node/308917 [accessed 2012-04-03] [WebCite Cache ID 66eOUUr8B]
14. Frog. 2010. Mobile Health: Project Masiluleke URL: http://www.frogdesign.com/work/project-m.html [accessed 2012-04-04] [WebCite Cache ID 66fVdB9z4]
15. aventh. Soul Beat Africa. 2011 Mar 21. Text To Change (TTC) URL: http://comminit.com/?q=africa/node/280980 [accessed 2012-04-03] [WebCite Cache ID 66eP1pUV1]
16. Tischler L. Fast Company. 2009. Welcoming guest blogger Robert Fabricant: designing for the unpredictable URL: http://www.fastcompany.com/blog/linda-tischler/design-times/welcoming-guest-blogger-robert-fabricant-designing-unpredictable [accessed 2012-04-03] [WebCite Cache ID 66eOwEZBr]
17. FHI360-SATELLIFE Health Information and Technology. 2010. ICT for health > empowering health workers to save lives: Uganda Health Information Network URL: http://www.healthnet.org/uhin [accessed 2012-04-03] [WebCite Cache ID 66eP82SwN]
18. Trucano M. EduTech. 2009 Apr 17. text2teach: what do we know about using mobile phones in education? URL: http://blogs.worldbank.org/edutech/category/tags/text2teach [accessed 2012-04-03] [WebCite Cache ID 66ePPyOcW]
19. Anantraman V, Mikkelsen T, Khilnani R, Kumar VS, Pentland A, Ohno-Machado L. Open source handheld-based EMR for paramedics working in rural areas. Proc AMIA Symp 2002:12-16 [FREE Full text] [Medline: 12463777]
20. SIMpill. 2010. The SIMpill Medication Adherence System is a potentially life-saving medication adherence solution that assists patients and/or carers in making sure that medication is taken as prescribed URL: http://www.simpill.com/ [accessed 2012-04-05] [WebCite Cache ID 66hWWeGLt]
21. Déglise C, Suggs LS, Odermatt P. Short message service (SMS) applications for disease prevention in developing countries. J Med Internet Res 2012;14(1):e3 [FREE Full text] [doi: 10.2196/jmir.1823] [Medline: 22262730]
22. Waruingi M, Underdahl L. AJFAND Online. 2009. Opportunity in delivery of health care over mobile devices in developing countries URL: http://www.ajfand.net/Volume9/No5/CommentaryWaruingi.pdf [accessed 2012-04-04] [WebCite Cache ID 66fql8MPL]
23. aventh. Soul Beat Africa. 2011 Mar 21. Mobile video for community health workers in Tanzania URL: http://www.comminit.com/?q=africa/node/320595 [accessed 2012-04-03] [WebCite Cache ID 66ePV7uOR]
24. RapidSMS. 2009. Malawi - Nutritional Surveillance URL: http://www.rapidsms.org/case-studies/malawi-nutritional-surviellence/ [accessed 2012-04-03] [WebCite Cache ID 66eQuIxq6]
25. Tumwesigye NM. Uganda Health Information Network. 2010. Cost effectiveness of personal digital assistants in health information system in Rakai and Lyantonde Districts, Uganda URL: http://web.idrc.ca/uploads/user-S/12759252461Cost_Effectiveness_UHIN.pdf [accessed 2012-04-04] [WebCite Cache ID 66fqGo4Ni]
26. Blaschke S, Bokenkamp K, Cosmaciuc R, Denby M, Hailu B, Short R. MobileActive.org. Brooklyn, NY: Mobile Development Solutions; 2009. Using Mobile Phones to Improve Child Nutrition Surveillance in Malawi: UNICEF Malawi and UNICEF Innovations URL: http://www.mobileactive.org/files/file_uploads/unicef_Malawi_CNS.pdf [accessed 2012-04-04] [WebCite Cache ID 66fR5I6GD]
27. Millennium Villages Project. ChildCount+. 2010. ChildCount+: A Community Health Events Reporting and Alerts System - A Concept Paper URL: http://childcount.org/files/ChildCountPlus_ConceptNote.pdf [accessed 2012-04-05] [WebCite Cache ID 66hkoFXTR]
28. The Economist. 2006 Mar 24. A text a day...the medical uses of mobile phones show they can be good for your health URL: http://www.economist.com/node/5655105 [accessed 2012-04-04] [WebCite Cache ID 66fpy5V6v]
29. Pop-Eleches C, Thirumurthy H, Habyarimana JP, Zivin JG, Goldstein MP, de Walque D, et al. Mobile phone technologies improve adherence to antiretroviral treatment in a resource-limited setting: a randomized controlled trial of text message reminders. AIDS 2011 Mar 27;25(6):825-834. [doi: 10.1097/QAD.0b013e32834380c1] [Medline: 21252632]

30. DKT International. 2012. DKT Mozambique URL: http://www.dktinternational.org/country-programs/mozambique/ [accessed 2012-07-13] [WebCite Cache ID 697wL3b6a]
31. McNab C. What social media offers to health professionals and citizens. Bull World Health Organ 2009 Aug;87(8):566 [FREE Full text] [Medline: 19704998]
32. Derenzi B, Borriello G, Jackson J, Kumar VS, Parikh TS, Virk P, et al. Mobile phone tools for field-based health care workers in low-income countries. Mt Sinai J Med 2011;78(3):406-418. [doi: 10.1002/msj.20256] [Medline: 21598267]
33. Zurovac D, Sudoi RK, Akhwale WS, Ndiritu M, Hamer DH, Rowe AK, et al. The effect of mobile phone text-message reminders on Kenyan health workers' adherence to malaria treatment guidelines: a cluster randomised trial. The Lancet 2011 Aug 27;378(9793):795-803 [FREE Full text] [doi: 10.1016/S0140-6736(11)60783-6] [Medline: 21820166]
34. Chang LW, Kagaayi J, Arem H, Nakigozi G, Ssempijja V, Serwadda D, et al. Impact of a mHealth intervention for peer health workers on AIDS care in rural Uganda: a mixed methods evaluation of a cluster-randomized trial. AIDS Behav 2011 Nov;15(8):1776-1784 [FREE Full text] [doi: 10.1007/s10461-011-9995-x] [Medline: 21739286]
35. Asiimwe C, Gelvin D, Lee E, Ben Amor Y, Quinto E, Katureebe C, et al. Use of an innovative, affordable, and open-source short message service-based tool to monitor malaria in remote areas of Uganda. Am J Trop Med Hyg 2011 Jul;85(1):26-33 [FREE Full text] [doi: 10.4269/ajtmh.2011.10-0528] [Medline: 21734120]
36. Datadyne. 2010. EpiSurveyor: Mobile Data Made Simple URL: http://www.datadyne.org/episurveyor [accessed 2012-04-03] [WebCite Cache ID 66ePdGAi6]
37. PATH. 2010. Interactive health information system to detect Japanese encephalitis in Asia URL: http://www.path.org/news/pr-041007.php [accessed 2012-04-03] [WebCite Cache ID 66ePkFjuQ]
38. Yang C, Yang J, Luo X, Gong P. Use of mobile phones in an emergency reporting system for infectious disease surveillance after the Sichuan earthquake in China. Bull World Health Organ 2009 Aug;87(8):619-623 [FREE Full text] [Medline: 19705013]
39. Bengtsson L, Lu X, Thorson A, Garfield R, von Schreeb J. Improved response to disasters and outbreaks by tracking population movements with mobile phone network data: a post-earthquake geospatial study in Haiti. PLoS Med 2011 Aug;8(8):e1001083 [FREE Full text] [doi: 10.1371/journal.pmed.1001083] [Medline: 21918643]
40. Malaria Consortium, World Health Organization. Moving Towards Malaria Elimination: Tools for Strengthening Malaria Surveillance in Cambodia. 2011. URL: http://www.malariaconsortium.org/userfiles/file/Resistance-Resources/Surveillence%20Tools%20-%20Moving%20Towards%20Malaria%20Elimination.pdf [accessed 2012-04-04] [WebCite Cache ID 66fXEZlgY]
41. Nariddh MC. Containment. 2011. Community malaria workers use SMS to report in real-time URL: http://malariacontainment.wordpress.com/category/cambodia/ [accessed 2012-04-03] [WebCite Cache ID 66eQ2K8Dv]
42. Healthy Child Uganda. 2010. Activities URL: http://healthychilduganda.org/healthy-child-uganda [accessed 2012-04-03] [WebCite Cache ID 66ePxLrzu]
43. Macueve G. e-Government Implementation in Mozambique: Transferring Lessons from the public sector. South African Computer Journal 2009;44:39-56.
44. Macanze J. Monitoring and Evaluation Report of PDAs for Malaria Monitoring in Maputo Province, Mozambique - Final Report. 2007. URL: http://web.idrc.ca/uploads/user-S/12053480321102534_ME_Report.pdf [accessed 2012-04-04] [WebCite Cache ID 66fWQO99x]
45. FHI360-SATELLIFE Health Information and Technology. 2010. ICT for health > empowering health workers to save lives: Mozambique Health Information Network URL: http://www.healthnet.org/mhin [accessed 2012-04-03] [WebCite Cache ID 66eQ7CfJ5]
46. Roll Back Malaria. 2010. SMS for Life: An RBM initiative URL: http://www.rbm.who.int/psm/smsWhatIsIt.html [accessed 2012-04-03] [WebCite Cache ID 66eQD363m]
47. WHO Global Observatory for eHealth. World Health Organization. 2006. Building Foundations for eHealth: Progress of Member States URL: http://www.who.int/goe/publications/bf_FINAL.pdf [accessed 2012-04-04] [WebCite Cache ID 66fYnKyZl]
48. SPIDER. 2010. ICT4MPOWER: ICT for Medical Community Empowerment URL: http://www.spidercenter.org/projects/ict4mpower-ict-medical-community-empowerment [accessed 2012-04-03] [WebCite Cache ID 66eQWbpTY]
49. Dimagi. 2010. CommCare - Community health mobile platform URL: http://www.commcarehq.org/home/ [accessed 2012-04-03] [WebCite Cache ID 66eQLvE7n]
50. Iyengar MS. Microsoft Case Studies. 2010. Smartphone-Based Medical Guides Aim to Improve Healthcare URL: http://www.microsoft.com/casestudies/-/Microsoft-Research/Smartphone-Based-Medical-Guides-Aim-to-Improve-Healthcare/4000004087 [accessed 2012-04-03] [WebCite Cache ID 66eQnvZoG]
51. CellScope. 2010. Mobile medical microscopy URL: http://cellscope.berkeley.edu/ [accessed 2012-04-04] [WebCite Cache ID 66fUYRNku]
52. Dimagi. 2010. E-IMCI is mobile device based job aid for implementers of IMCI URL: http://www.dimagi.com/mobile-e-imci/ [accessed 2012-04-03] [WebCite Cache ID 66eQf8ef5]
53. Iyengar MS. slideshare. 2009. The GuideView mHealth System URL: http://www.slideshare.net/gueste312b0/the-guideview-mhealth-system [accessed 2012-04-03] [WebCite Cache ID 66eR4vQOL]

54. Bocock P. Global Health Impact: Management Sciences for Health. 2010. Looking for mHealth Sustainability? Look for Country Ownership and See how Africa is Leading the Way URL: http://blog.msh.org/2010/12/07/looking-for-mhealth-sustainability-look-for-country-ownership-and-see-how-africa-is-leading-the-way/ [accessed 2012-04-05] [WebCite Cache ID 66hgr2gdc]
55. Purdy C. HealthWorks Collective. 2012. The Developing World Uses Social Media to Promote Safe Sex URL: http://healthworkscollective.com/psalber/32330/can-internet-help-you-slip-condom-brazil-or-mozambique [accessed 2012-07-13] [WebCite Cache ID 697wZCQ4Z]
56. Sharp JW. eHealth. 2011. Social media in health care: What's Working, What's Not URL: http://ehealth.johnwsharp.com/2011/04/14/appropriate-use-of-social-media-in-healthcare-organizations/ [accessed 2012-07-13] [WebCite Cache ID 697seHwfF]
57. World Health Organization, UNICEF. Unicef. 2012. WHO/UNICEF Joint Statement: Integrated Community Case Management (iCCM) URL: http://www.unicef.org/health/files/iCCM_Joint_Statement_2012.pdf [accessed 2012-07-08] [WebCite Cache ID 690mjoIp2]
58. Braa J, Macome E, Mavimbe J, Nhampossa L, Cossa J, Jose B, et al. The Electronic Journal of Information Systems in Developing Countries. 2001. A study of the actual and potential usage of information and communication technologies at district and provincial levels in Mozambique with a focus on the health sector URL: http://www.ejisdc.org/ojs2/index.php/ejisdc/article/view/27 [accessed 2012-04-04] [WebCite Cache ID 66fUIldzz]
59. Kimaro H, Nhampossa L. Analyzing the problem of unsustainable health information systems in less?developed economies: case studies from Tanzania and Mozambique. Information Technology for Development 2005;11:273-298. [doi: 10.1002/itdj.20016]

Abbreviations

AESSIMS: Acute Encephalitis Syndrome Surveillance Information System
FIND: Foundation for Innovative New Diagnostics
GPRS: general packet radio service
GPS: Global Positioning System
ICCM: integrated community case management
MHIN: Mozambique Health Information Network
PCM: please call me
PDA: personal digital assistant
SIM: subscriber identity module
SMS: short message service
TTC: Text-to-Change
UHIN: Uganda Health Information Network
USSD: Unstructured Supplementary Service Data
VoIP: voice over IP
WHO: World Health Organization

Edited by G Eysenbach; submitted 07.04.12; peer-reviewed by E Afari-kumah, A Fuad, S Langrial; comments to author 21.05.12; revised version received 19.10.12; accepted 08.11.12; published 25.01.13.

Please cite as:
Källander K, Tibenderana JK, Akpogheneta OJ, Strachan DL, Hill Z, ten Asbroek AHA, Conteh L, Kirkwood BR, Meek SR
Mobile Health (mHealth) Approaches and Lessons for Increased Performance and Retention of Community Health Workers in Low- and Middle-Income Countries: A Review
J Med Internet Res 2013;15(1):e17
URL: http://www.jmir.org/2013/1/e17/
doi:10.2196/jmir.2130
PMID:23353680

Original Paper

Quantifying Short-Term Dynamics of Parkinson’s Disease Using Self-Reported Symptom Data From an Internet Social Network

Max Little[1], B.Sc (Hons), PhD; Paul Wicks[2], PhD; Timothy Vaughan[2], PhD; Alex Pentland[1], PhD

[1]Human Dynamics Group, Media Lab, Massachusetts Institute of Technology, Cambridge, MA, United States

[2]Research and Development, PatientsLikeMe Inc., Cambridge, MA, United States

Corresponding Author:
Max Little, B.Sc (Hons), PhD
Human Dynamics Group
Media Lab
Massachusetts Institute of Technology
77 Massachusetts Avenue
Cambridge, MA, 02139
United States
Phone: 1 617 253 1000
Fax: 1 617 253 1000
Email: maxl@mit.edu

Abstract

Background: Parkinson’s disease (PD) is an incurable neurological disease with approximately 0.3% prevalence. The hallmark symptom is gradual movement deterioration. Current scientific consensus about disease progression holds that symptoms will worsen smoothly over time unless treated. Accurate information about symptom dynamics is of critical importance to patients, caregivers, and the scientific community for the design of new treatments, clinical decision making, and individual disease management. Long-term studies characterize the typical time course of the disease as an early linear progression gradually reaching a plateau in later stages. However, symptom dynamics over durations of days to weeks remains unquantified. Currently, there is a scarcity of objective clinical information about symptom dynamics at intervals shorter than 3 months stretching over several years, but Internet-based patient self-report platforms may change this.

Objective: To assess the clinical value of online self-reported PD symptom data recorded by users of the health-focused Internet social research platform PatientsLikeMe (PLM), in which patients quantify their symptoms on a regular basis on a subset of the Unified Parkinson’s Disease Ratings Scale (UPDRS). By analyzing this data, we aim for a scientific window on the nature of symptom dynamics for assessment intervals shorter than 3 months over durations of several years.

Methods: Online self-reported data was validated against the gold standard Parkinson’s Disease Data and Organizing Center (PD-DOC) database, containing clinical symptom data at intervals greater than 3 months. The data were compared visually using quantile-quantile plots, and numerically using the Kolmogorov-Smirnov test. By using a simple piecewise linear trend estimation algorithm, the PLM data was smoothed to separate random fluctuations from continuous symptom dynamics. Subtracting the trends from the original data revealed random fluctuations in symptom severity. The average magnitude of fluctuations versus time since diagnosis was modeled by using a gamma generalized linear model.

Results: Distributions of ages at diagnosis and UPDRS in the PLM and PD-DOC databases were broadly consistent. The PLM patients were systematically younger than the PD-DOC patients and showed increased symptom severity in the PD off state. The average fluctuation in symptoms (UPDRS Parts I and II) was 2.6 points at the time of diagnosis, rising to 5.9 points 16 years after diagnosis. This fluctuation exceeds the estimated minimal and moderate clinically important differences, respectively. Not all patients conformed to the current clinical picture of gradual, smooth changes: many patients had regimes where symptom severity varied in an unpredictable manner, or underwent large rapid changes in an otherwise more stable progression.

Conclusions: This information about short-term PD symptom dynamics contributes new scientific understanding about the disease progression, currently very costly to obtain without self-administered Internet-based reporting. This understanding should have implications for the optimization of clinical trials into new treatments and for the choice of treatment decision timescales.

(J Med Internet Res 2013;15(1):e20) doi:10.2196/jmir.2112

KEYWORDS

Parkinson’s disease; social networks; medical informatics; symptoms; pharmacodynamics

Introduction

Parkinson’s disease (PD) is a relatively common, progressive neurological disorder affecting approximately 0.3% of the general population in industrialized countries [1]. It generally affects people over 60 years, but rarely younger people under the age of 40 years also develop the disease. PD is considered a movement disorder (ie, it affects the ability to perform normal voluntary motion), but patients also experience cognitive impairment and emotional/mood disturbances. The classic movement symptoms of PD include exaggerated tremor, rigidity, and slow or hesitant motion. These movement problems often have a substantial negative impact on the ability of the patient to perform essential everyday activities, such as bathing, dressing, turning in bed, walking unaided, and getting up from a sitting position. The cause of PD is currently thought to be the loss of dopaminergic neurons in an area of the brain known as the substantia nigra. PD is incurable, and there are no absolutely conclusive diagnostic tests. The most accurate diagnosis based on behavioral symptoms achieves, at best, 90% accuracy when compared to postmortem pathological examination [2].

The mortality rate of patients with the disease is significantly increased relative to healthy people [1]. There are a few approaches to treating the symptoms of PD. The first line of defense is the drug levodopa that replenishes dopamine in the substantia nigra, which reduces the severity of movement symptoms. However, this drug tends to become less effective over time and can also lead to severe side effects, such as involuntary movements (dyskinesias). Surgical treatments, such as deep brain stimulation, have been shown to be effective for many patients who do not respond or have ceased to respond to drug treatments. Current scientific understanding holds that the severity of PD symptoms will smoothly increase over time, faster at first and often leveling out in the later stages [3].

Trials for new treatments and assessing the effectiveness of treatments require objective data about symptom severity. A coarse quantitative measure of symptom severity is the Hoehn and Yahr (HY) ordinal scale [4] that assigns a number from 0 to 5, with 0 being healthy and 5 denoting severe disability. This has been largely supplanted by the ordinal Unified Parkinson’s Disease Rating Scale (UPDRS) (version 3.0) [5] and associated tests [6], which are more time consuming and expensive to administer, but are more precise. The most commonly used parts of the UPDRS (Parts I, II, and III) range on a scale from 0 (healthy) to 176 (severe disability) [5], although a simple and accurate formula exists to predict HY from UPDRS [7].

The UPDRS values have been collected for patients at all stages of the disease, and there is substantial research data available on PD symptom progression quantified on this scale. This kind of data has been used to calibrate models of PD symptom progression over the course of years to decades [3]. However, the full UPDRS is a complex test that requires expertise to administer (even if that expertise can be taught to general medical personnel [5]), attendance of the patient in the clinic, and the average time for administration of the full test is approximately 17 minutes [8]. Unfortunately, these difficulties mean that it is usually prohibitive to objectively score PD symptom severity on timescales shorter than 3 months (low frequency). Since most longitudinal UPDRS data is low frequency, objective information about symptom dynamics occurring on a shorter timescale than 3 months (high-frequency data) is lacking.

There are many clinical situations in which high-frequency symptom dynamics would be useful. For example, in testing new drug treatments there is a trade-off between minimizing exposure to the novel drug to reduce the risk of unknown side effects and maximizing the opportunity to detect significant changes in symptoms. This temporal trade-off cannot be optimized on a quantitative basis without high-frequency data upon which to base the statistical analysis. Similar issues arise in diagnosis where PD is suspected. If, in conjunction with movement symptoms on 1 side of the body only, taking levodopa leads to a reduction in symptom severity, the patient is highly likely to have PD [2]. However, there is still a non-negligible chance that the patient has some other neurological disorder with PD-like symptoms, such as progressive supranuclear palsy. This disease can progress very rapidly, so it is important to diagnose this quickly. Thus, the window of this “exploratory” prescription of levodopa for differential diagnosis must be made as short as possible. However, it should not be so short that rapid, natural fluctuations in symptoms confound proper diagnosis.

Recently, health-focused Internet websites have been established that allow users to track their disease progression by using surveys and other remote monitoring devices, for example. We obtained the UPDRS data from PD users of the PatientsLikeMe website [9], which has recruited over 6000 PD patients worldwide since 2007. Some of these patients are particularly dedicated diarists who have documented their symptoms on a regular basis over a number of years. The result is an unprecedented, high-frequency symptom dataset that has the potential to be used to address some of the shortcomings of existing low-frequency clinical data. For example, if the data are sufficiently accurate, it could be used to supplement in-clinic checkups between visits. Similar data were used for another neurological disease (amyotrophic lateral sclerosis) to refute the idea that lithium carbonate slowed the progression of that disease [10]. The purpose of this study is an exploratory investigation into the high-frequency dynamics and other properties of this novel PD dataset to assess the clinical value of these data.

Methods

Patient Recruitment and Data Collection

The main outcomes of this study were quantified by using the UPDRS. This scale consists of 5 parts: Part I covers cognitive, behavioral, and mood symptoms; Part II evaluates activities of daily living; Part III measures motor symptom severity; and Parts IV and V contain HY stage and an evaluation of daily living activities on the Schwab and England scale [6]. Parts I to III contain separate sections, each with a score ranging from 0 (no symptoms) to 4 (severe symptoms). Part I has 4 sections, and Part II has 13 sections.

Two data sources were used: the PatientsLikeMe (PLM) dataset and the Parkinson’s Disease Data and Organizing Center (PD-DOC) dataset [11]. The PLM data were used to provide long-term quantification of individual symptoms occurring on a timescale shorter than 3 months. The data are entirely self-reported. Users sign up to the website where they can enter demographic details, information about their disease course and symptoms, and their treatment history. Specifically, we collected age, gender, treatment status, HY staging, and UPDRS Parts I (mentation, behavior, and mood) and II (activities of daily living) information. Part III of the UPDRS (motor symptoms) was excluded because the collection of this data was deemed not suitable for self-report. Not all self-reported symptoms were accompanied by treatment status indications.

At the time of preparation of this manuscript, the PLM dataset contained 6074 PD patients, of which 2931 completed at least 1 UPDRS survey and entered their birth date and date of diagnosis. Patients were included in this study if they reported at least 15 UPDRS scores with a maximum average UPDRS reporting interval of 65 days between reports. This led to 100 patients being included in this study, and a mean of 29 (SD 14) symptom self-reports per patient (total of 2896 reports), with a reporting interval mean of 45 days (SD 12). The mean age of the selected patients was 54 years (SD 9) at diagnosis, of which 52 were female, 48 male. Patients began self-reporting symptoms approximately 1 year after diagnosis, on average. The total time interval covered by self-reporting per patient, from the first report to the last, was 3.1 years (SD 0.8), and all reports were prospective (after date of joining the website). Patients contributing to the PLM data agreed to the terms and conditions of the website when they enrolled, which included granting permission to PLM to use their medical data for research purposes [12]. Qualitatively, the PLM dataset represents a large number of frequent Part I and II UPDRS reports and treatment status across a medium-size cohort of young to middle-aged patients.

The PD-DOC dataset contains data on PD patients from multiple clinical centers in the United States across several trials with data collected by clinicians over the period from 2006 to 2011 to aid the process of statistical analysis of PD and for the design and planning of clinical trials into treatments. In this study, it was used as a reference dataset to verify the PLM data and to provide background data on PD. Data collection was coordinated by the University of Rochester, Rochester, NY. The set represents UPDRS symptom reports from 564 individuals with PD, of which 200 were female and 364 male, and a mean age of 59 years (SD 10) at diagnosis. In PD, during the day there will be “on” periods when the symptoms of PD are suppressed, to a greater or lesser degree, by the treatment, and “off” periods when the full symptoms reoccur even while taking treatment. In the “on” state, 1612 UPDRS scores were recorded and 354 were recorded in the “off” state. There were a mean of 2.9 (SD 0.9) symptom reports, covering an average of 1.9 years (SD 0.9) per patient. Ethical approval was obtained from the independent review boards of each US medical center contributing patient details to the dataset. In contrast to the PLM dataset, PD-DOC can be described as data from a large number of middle-aged to older patients with clinical UPDRS reports collected on an infrequent basis.

Validating the PatientsLikeMe Dataset

At the outset, the concept of symptom self-reporting may raise data reliability questions, primarily because it could be suspected that untrained nonclinical raters may be more prone to certain systematic errors or biases than trained clinical raters. For example, they may tend to be biased toward repeating previous measurements, or may have more inconsistent interpretations of specific questions across tests than trained clinical staff. Previous research has shown that when PD patients without dementia self-report UPDRS Parts I and II scores, the scores are consistent with those assessed by the neurologist assessing them [13]. To our knowledge, there have been no similar assessments into the reliability of self-reported UPDRS Parts I and II scoring conducted online under nonclinical circumstances.

To address this issue, we compared the PLM dataset against the PD-DOC data that we considered to be a gold standard clinical reference set. The distributions of UPDRS Parts I and II values and ages at diagnosis were compared visually on quantile-quantile (q-q) plots: if the distributions were of the same form (ie, the same up to a transformation of location and scale, typically the mean and standard deviation), then on the q-q plot the data will lie, approximately, on a straight line [14]. In addition, if the location and scale parameters are the same, the data will lie on a line with a slope equal to 1. Numerical comparisons were made by using the 2-sample Kolmogorov-Smirnov (K-S) test applied to the *z*-scored data (ie, data in which the mean was removed and then divided by the standard deviation). This high-precision test was applied to quantify the results of the visual q-q plot analysis.

Trend Estimation

To analyze the dynamics of PD over short time periods, it is necessary to remove the effect of trends that occurred due to the natural progression of the disease over that timescale. One widespread approach to modeling disease progression is the use of hierarchical mixed-effects models [15]. These are commonly applied in pharmacodynamics studies [16]. Considerable effort over the preceding decades has increased the sophistication of these models from their origins in simple linear mixed-effects models by incorporating additional features, such as smooth [15] or abrupt nonlinearities in progression [3], nonparametric progression curves [17], and more recently, clustering of individuals into arbitrary groupings using nonparametric Bayes techniques [18]. In PD, pharmacodynamic studies have fitted the smooth, Gompertz sigmoidal curve as a model for progression over the lifetime of the patient with parameters estimated on low-frequency data [3].

A predominant feature of these models is pooling data between subjects. Drawing on specific knowledge about underlying physiological processes (eg, in virology, the mechanism of viral infection of cell populations can give a biologically plausible functional form for the curve). Then the problem becomes one of estimating the parameters for the curve, also known as a regression problem. When there is insufficient progression data about each individual to get reliable (low variance) individual parameter estimates, a global model that fits the data pooled over all individuals can be more reliable, but biased with respect to each individual. By assuming that the individual regression parameters are random variables, it is possible to form compromise parameter estimates by using an appropriate mix of the global and individual models: this is the main premise of (2-level) hierarchical modeling.

In our case, we wished to perform an exploratory smoothing of the PLM data that made use of as few assumptions as possible, and had easily traceable logic from underlying assumptions to the results obtained by a simple statistical inference procedure. Also, because we had adequate data at the individual level, we did not need a pooled model. These considerations meant that existing mixed-effects models were not suited to our application: they require complex inference schemes that involve approximations (because nonlinear models are generally analytically intractable) that obscure the interpretation of the results, and would be biased from the perspective of the individual [15-17].

We used a piecewise linear convex regression smoothing approach (see Multimedia Appendix 1), which can approximate smooth, nonlinear progression as a series of lines, and can also naturally model abrupt changes in progression. The only assumption about the resulting curve is that it has minimal total absolute curvature (second derivative against time) given a fixed total mean squared error with respect to the individual's PLM data. Note that this model is related to, but much simpler than, the nonparametric spline mixed-effects model of Rice and Wu [17]. In contrast to the Rice and Wu model, however, the inference problem is convex (it has a verifiable optimum solution), which is solved by stable computations whose convergence properties are guaranteed [19].

Residual Modeling

The trend identified above is subtracted from the UPDRS data to obtain the residuals. Modeling these fluctuations allows us to quantify the high-frequency dynamics of PD symptoms. Trends in the size of these fluctuating residuals can be detected by using a variety of methods, but because of the specifics of the trend estimation algorithm described previously, we modeled the size of the residuals against time since diagnosis by using a gamma generalized linear model (see Multimedia Appendix 1).

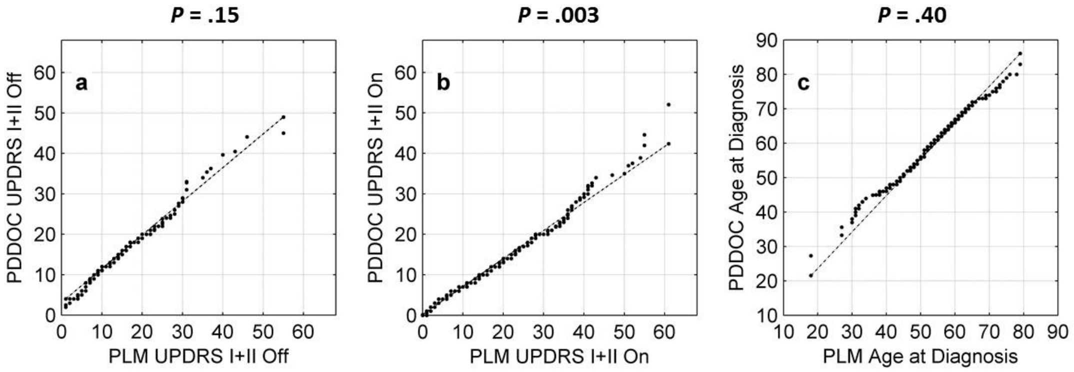
All analyses were carried out using specialized software written for the MATLAB platform version R2007a (MathWorks Inc, Natick, MA, USA). Creative Commons-licensed trend estimation software is distributed with this publication.

Results

PatientsLikeMe Dataset Validation

The PLM and PD-DOC datasets agree in terms of the broad shape of the distribution of UPDRS values and ages at diagnosis (Figure 1). The K-S test results indicate that, up to a change in standard deviation and mean, "off" UPDRS values and ages appear to come from the same distribution, whereas "on" UPDRS values do not (Note that for the K-S test, when *P*<.05, the null hypothesis that the *z*-scored data come from the same distribution can be rejected at the 95% level). Therefore, there are some systematic differences (see discussion section), but the fact that the PLM and PD-DOC distributions are broadly similar in distribution is good evidence that the online PLM dataset is as reliable as objective clinical data about patient's symptom severity.

Figure 1. Validating the online self-reported PatientsLikeMe (PLM) dataset against the clinically scored Parkinson's Disease Data and Organizing Center (PD-DOC) reference dataset. Visual comparisons using quantile-quantile plots; statistical comparisons using the 2-sample Kolmogorov-Smirnov (K-S) test applied to z-scored data (K-S test results displayed as the P values above graphs). (a) Unified Parkinson's Disease Rating Scale (UPDRS) values (sum of Parts I and II) for values labeled as "off" treatment in the PLM dataset against values labeled as "off state" in the PD-DOC data; (b) as with (a), except for the "on" treatment/state labels; (c) ages at diagnosis.



Trend Estimation

After performing trend estimation (Figure 2 illustrates the selection of the regularization constant and the resulting trend), our next finding is that although most patients do have smooth progression in symptom severity over time with small to moderate short-term variability (Figure 3), there are an interesting and important minority who do not (Figure 4). In the former group, we found patients with very predictable increases in symptom severity, increases that slow over time (Figure 3a, c, and d). We also see patients responding well to treatment with gradually decreasing symptom severity that eventually reaches a plateau (Figure 3b). These patients all conform to the current consensus picture of smooth, long-term symptom changes (eg, following the smooth Gompertz curve [3]). However, in the nonconforming group, we found evidence for unpredictable medium-term changes (Figure 4a and b), and occasional rapid increases (outliers) in otherwise smooth progression (Figure 4c and d).

Figure 2. Trend fitting and residual modeling of the self-reported Unified Parkinson’s Disease Rating Scale (UPDRS) values (Parts I and II) in the PatientsLikeMe (PLM) dataset. (a) Absolute values of residuals obtained by subtracting the long- to medium-term trend from the raw values (natural logarithmic vertical scale), plotted against time since diagnosis in years. The blue line (formula inset) shows the estimated most-likely relationship between time since diagnosis and average absolute residual value. The gray lines are the 95% CI for the relationship; (b) the relationship between average absolute residual and time since diagnosis in (a) shown on a linear vertical scale; (c) UPDRS trend, used to calculate residuals, estimated from an example patient; (d) choice of trend regularization constant for (c), occurring at the smallest value of the cross-validated trend test error.

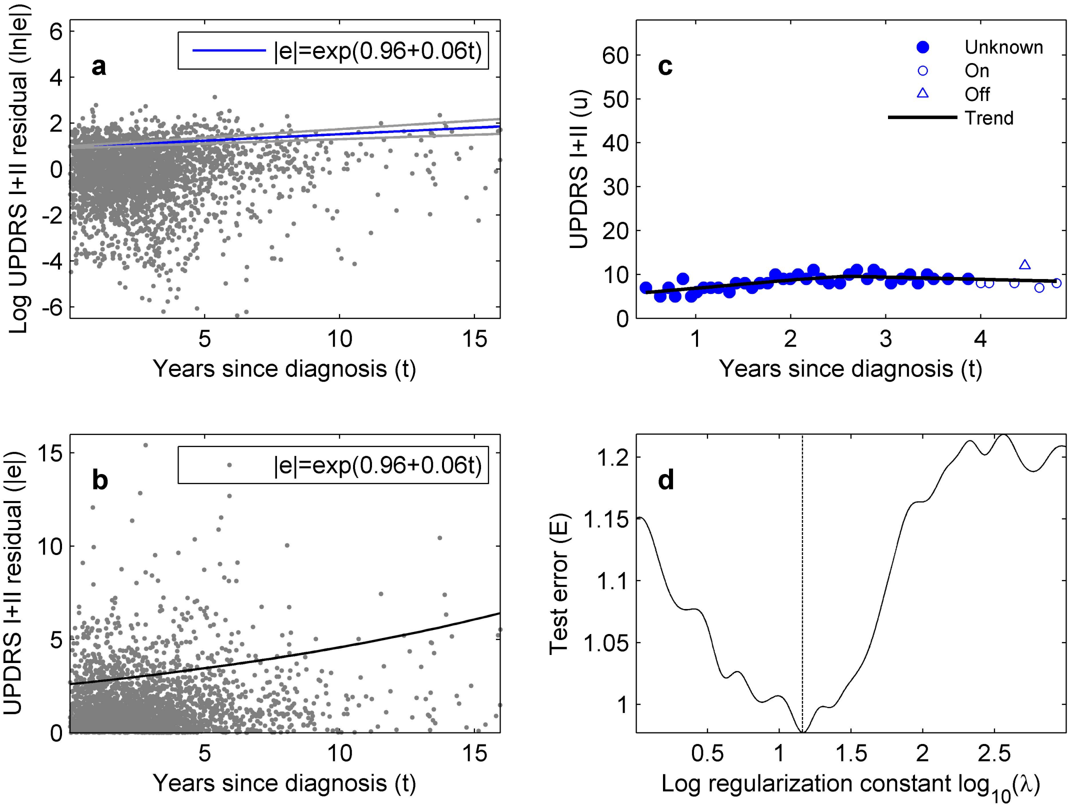


Figure 3. Four examples of patients in the PatientsLikeMe (PLM) dataset whose self-reported symptom dynamics conform to the current consensus picture of slow, predictable Parkinson's disease symptom progression. Increase is generally smooth, variation around the trend (residuals) are generally small. "Unknown" refers to data in which the patient did not state whether they were on treatment (on) or off treatment (off) at the time of the symptom report.

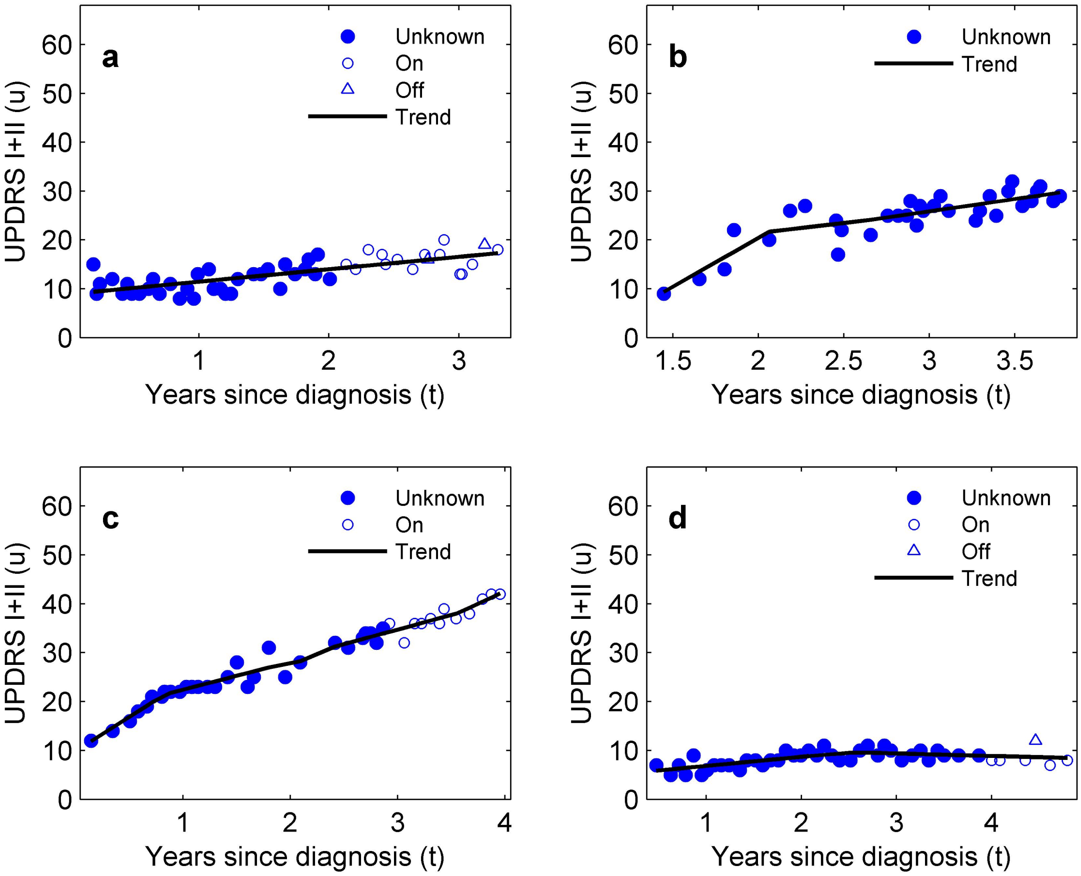
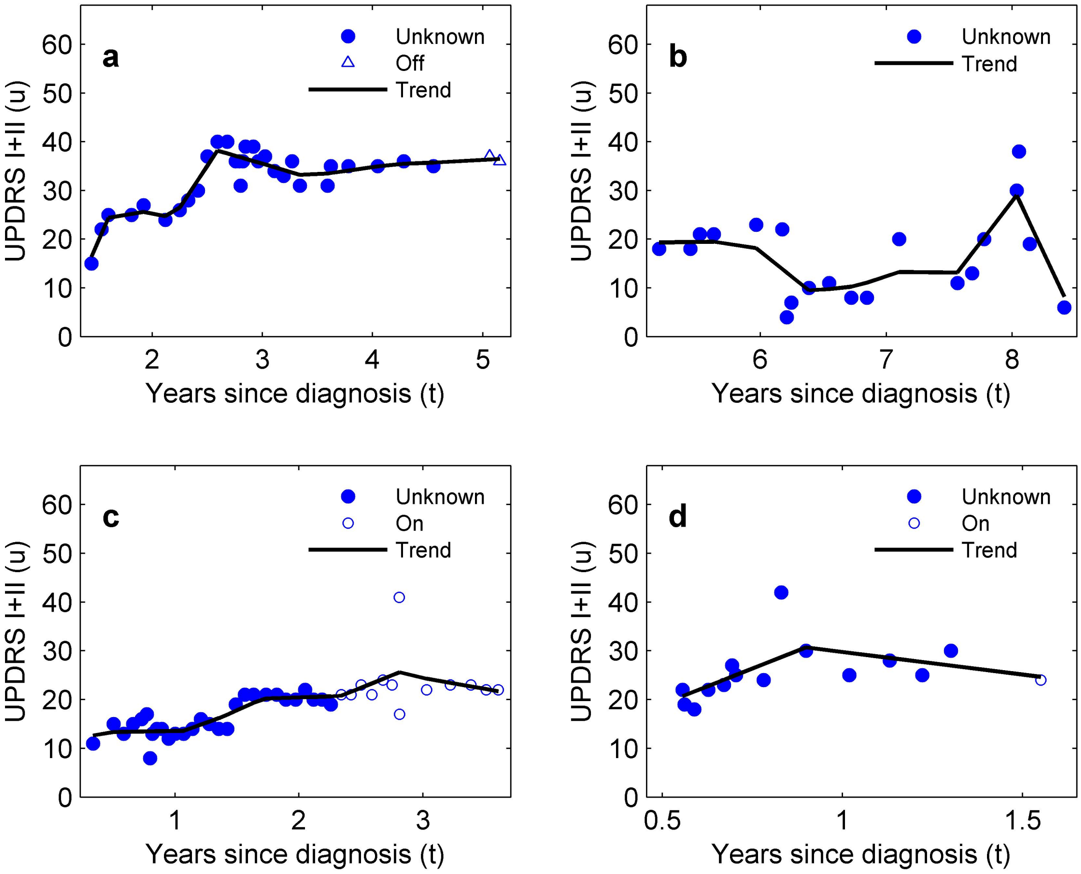


Figure 4. Four examples of patients in the PatientsLikeMe (PLM) dataset whose self-reported symptom dynamics diverge from the current consensus picture of smooth, gradual Parkinson's disease progression. (a,b) Large, abrupt changes in the trend of symptom severity occurring over periods of a few weeks or months; (c,d) examples of single, large deviations from the trend in otherwise smooth progression. "Unknown" refers to data in which the patient did not state whether they were on treatment (on) or off treatment (off) at the time of the symptom report.



Residual Modeling

Residuals quantifying short-term fluctuations in symptom severity on the scale of days to weeks (which affect all patients to a lessor or greater extent) increase steadily in amplitude with time since diagnosis (Figure 2a and b). At diagnosis, average symptom severity variation is 2.6 points, rising to 5.9 points 16 years later. This finding is not simply a systematic consequence of the nonnegative UPDRS scale: if the short-term variation is to be symmetric about the long-term trend, then as the score becomes small, the residuals must get smaller to avoid negative UPDRS values. However, we find that the residuals are significantly positively skewed (*t* test against zero skewness rejected with *P*<.001, based on 1000 replicate bootstrapped skewness values). A similar argument would hold for very large UPDRS values because the scale has a maximum value of 68. The PLM data does not contain sufficient information about severely disabled patients with very high UPDRS Parts I and II values (less than 10% of the symptom reports in the database have stage 4 or 5 HY), and so this argument cannot be tested with the data available to this study.

Discussion

Summary of Results

This study addressed the topic of quantifying trends and variability in PD symptoms that occurs on a timescale shorter than 3 months. A dataset of 100 PD patients with symptoms, self-reported on a standard clinical scale, was analyzed. Although we have found examples of specialized studies collecting weekly UPDRS values in the literature (eg, Goetz et al [20] recording motor UPDRS weekly for 8 weeks from 16 patients to assess the effect of switching dopamine agonists), to our knowledge, this rapid self-reporting of PD symptoms stretching over many years in the PLM data is unprecedented among existing reference clinical datasets. With appropriate feedback and social network community engagement, this dataset has the potential to grow quickly at little marginal cost per patient because the usefulness of the network grows with the square of the network size (an observation known as Metcalfe's law).

Validation demonstrated the high-frequency self-reported data are consistent with a low-frequency clinical dataset in common use in clinical PD studies. The distributions of PLM to PD-DOC off scores are essentially the same (Figure 1a). One systematic difference is that the mean age of PLM patients is approximately 6 years younger than the mean age of PD-DOC patients (Figure 1c) which is most likely a sociological effect: younger patients are generally more technologically aware, able, or willing to share their personal data. Similar patterns have been identified in other conditions, such as multiple sclerosis [21], in which the average online patient was 4 years younger than patients in a comparable clinical reference dataset.

Another systematic difference is that the symptom scores for PLM patients labeled as on treatment are biased upwards by comparison to the PD-DOC data (Figure 1b). Furthermore, the largest symptom scores (>40 UPDRS points) for PLM patients are much more common than in the PD-DOC dataset (this is the reason why, even after *z*-scoring, the K-S test fails). The most plausible explanation for this difference in on scores (PD-DOC) versus on treatment (PLM) is because of differences in interpretation of the meaning of "on."

As discussed earlier, and repeated here for clarity, during the day there will be "on" periods when the symptoms of PD are suppressed, to a greater or lesser degree, by the treatment, and "off" periods when the full symptoms reoccur, even while taking treatment. The on/off terminology, therefore, has this somewhat specialized clinical meaning. In the PLM dataset, when completing UPDRS self-reports, patients are presented with the following question: "When you answer these questions, are you thinking about how you are on treatment or off treatment?" and they can respond by selecting either "on treatment" or "off treatment." In the PD-DOC dataset, which is collected by trained clinical staff, it can be assumed that the on/off terminology is used according to the clinically accepted definition described previously. By contrast, with self-reporting in the PLM dataset, it is more likely that the on/off labels refer to taking treatment (on) versus not taking treatment (off), and it is unclear whether patients are generally aware of the accepted clinical meaning of the on/off terms. PLM self-reporters can indicate that this is their UPDRS value while on treatment, and this would partly concur with the clinical on state. Similarly, PLM self-reports indicating off treatment might, only partly, overlap with the clinical off state in the PD-DOC dataset.

It is likely that the PLM on label includes many scores that would be considered, clinically, as off instead, because they refer to unsuppressed symptoms occurring while the patient is actively taking treatments (the clinical off condition). This would lead to the increased scores we observed.

To identify trends in symptom progression, a cross-validated, convex piecewise linear smoothing technique was applied to the self-reported data. After subtracting the trend from the self-reported scores, the remaining residual variations appeared to increase with time since diagnosis. Furthermore, a minority of patients were shown to deviate quite considerably from the existing consensus understanding that proposes smooth, gradual change in symptoms over time. Our conclusion is that these residuals are naturally heteroscedastic.

The variations in symptom severity we detected are unlikely to be clinically irrelevant fluctuations; previous studies have estimated the minimal clinically important difference (CID) in total UPDRS (Parts I-III) values as approximately 4.1 to 4.5 points [22]. The maximum value of measuring only Parts I and II is 68, whereas the total UPDRS value is 176 points. From this, we can get a rough estimate of 1.7 (calculated as $68/176 \times 4.3$) as the minimal CID for the data in this study, which implies that at the time of diagnosis, the average residual variation of 2.6 around the trend that we found is larger than the minimal variation in symptoms needed to trigger clinical decisions. Later, at 16 years after diagnosis, the same calculation shows a moderate CID of 3.3 points, so the average variation we find here (5.9 points) could be very misleading if taken out of context (eg, in a clinical trial for a new drug treatment).

Limitations

This study collected self-reported data about cognitive, behavioral, mood symptoms, and impairment in activities of daily living. PD is primarily a movement disorder; therefore, it is important to also be able to quantify movement symptoms. Nonetheless, activities of daily living are significantly impaired by motor deterioration, so this section of the UPDRS measures motor symptoms indirectly. Because the UPDRS is additive, even if motor symptoms increase smoothly in severity according to a long-term trend, the total UPDRS score (Parts I-III) would still show the effect of variability in Parts I and II that we observe here. Previous low-frequency studies showed evidence of the kind of variability that we found here in motor symptoms, such as bradykinesia, rigidity, and tremor [3]. Thus, we have some confidence that the explicit inclusion of a direct quantification of motor symptoms, although an important addition that would alter our assessment of the specific numerical results presented here, will not fundamentally alter our conclusions.

The PLM website has no mechanism to require patients to return to the site and enter new symptom reports, but it is possible that many patients only return to the site to enter their symptoms when they have experienced a symptom fluctuation. However, patients are unlikely to agree on what level of change in UPDRS constitutes a reportable fluctuation; and so, we would expect to see fluctuations of all sizes and differing reporting intervals in the dataset. Also, we would not expect to find regular time intervals between reports (if symptom fluctuations are indeed random). Therefore, there is no reason to believe that such fluctuation-triggered reporting is a significant source of bias in our results.

The standard CID calculations in UPDRS are performed on cross-sectional data, and refer to the symptom variation around the average across all individuals in the PD population [22]. Therefore, to draw meaningful comparisons against this literature we have performed the equivalent pooling across all individuals. These CID calculations make the statistical assumption that the patients all come from a homogenous group sharing the same UPDRS distribution. Our findings here probably indicate that this assumption may not be statistically accurate because we have found quite significant differences in symptom progression and magnitude of variation. Further statistical analysis may be needed to identify the nature of any systematic differences or subgroups in residual distribution.

Implications for Parkinson’s Disease Research and Clinical Practice

We detected fundamental variability in symptoms on timescales less than 3 months that all patients at all stages of the disease seem to show. We note that the variability captured by the residuals we see here is not the same as the variability usually associated with fluctuators, the clinical term used to refer to patients with severe symptoms, usually in the later stages of the disease, who experience intermittent responsiveness to drug treatments [23].

Typical of many eHealth studies [24], we found a large attrition rate. Of the more than 6000 PD patients registered on the site, the fraction of sufficiently committed users is small (less than 2%). It can be estimated that entering 30 symptom reports would require, on average, approximately 7.5 hours of patients’ time in total, using timing information derived from self-administered paper data entry [5]. This is a lot of time to dedicate to entering data into a computer if there is no obvious reward (eg, financial compensation frequently used with clinical trials), even if spread over nearly 3 years, and is one plausible explanation for this severe attrition rate. It is possible that patients who are this dedicated are a select group who may introduce some, as yet unknown, bias into the results. Nonetheless, aside from this group being younger than typical clinical populations, we are not aware of any particular reason why the results we present here would be biased by focusing on a core of more dedicated symptom diarists.

We found that altering the inclusion/exclusion criteria from the PLM dataset did not lead to significant changes in the residual model.

The fundamental variability we detected here represents a critical factor in clinical decision making: knowing what sort of variability to expect is important because it determines how long to wait to detect a significant improvement in symptoms following a change in treatment regime, for example. The explicit information provided here could also be used to build improved progression models, for example, knowledge of the distribution of the residuals can be used to derive more accurate statistical model-fitting algorithms.

It is difficult to speculate on the origins of such heterogeneity in progression, but other studies have identified different clinical subtypes of PD [25]. It is possible that this might also be reflected in different progression profiles. Future research using this kind of high-frequency data might be able to identify different progression subtypes.

The main issue we identified with existing clinical PD symptom data is that it is an undersampling of the high-frequency data we presented here [26]; that is, because the sampling frequency is so low, it does not adequately represent the kind of symptom fluctuations that most patients experience on timescales shorter than 3 months.

The existence of such nonconforming patients is of critical importance to trial design in which it is typically assumed, based on current understanding of PD symptom progression, that symptoms will change slowly over the duration of the trial. However, this is not always true (Figure 4a and b). Recruiting patients into trials with the expectation that symptoms will change slowly over that period may lead to questionable results, including the failure of trial statistics to show clinically significant outcomes, not as a consequence of the failure of the treatment under test, but because of a failure to incorporate such nonconforming progression into the statistical procedures used to analyze the data.

We see this study as a prelude to the next logical step of increasing the frequency of objective symptom measurement even further. For example, we envisage these results being of utility in the design of novel, noninvasive, objective symptom severity quantification algorithms. Methods based on voice [27] or accelerometry [28], particularly using smartphones, seem promising because they offer the potential to track the effectiveness of choices in drug dosage and timing in real time. These new methods will require high-frequency reference symptom data for verification and current clinical reference data, such as the PD-DOC database, are insufficiently detailed for this purpose.

The ability to remotely self-administer tests for PD symptom severity data offers considerable cost reductions for most clinical applications by reducing the cost of clinical staff time and transport for patients during routine checkups, and lowering the costs of recruitment and tracking of patients in clinical trials, for example. Finally, there is the potential to use this kind of high-frequency data to fit models that can be used for prognostics to predict each patient’s future symptom severity.

Acknowledgments

Max Little is funded by a Wellcome Trust-MIT postdoctoral fellowship, grant number WT090651MF.

Conflicts of Interest

PW and TV are employees of PatientsLikeMe and own stock options in the company. The PatientsLikeMe research and development team has received research funding from Abbott, Accorda, Avanir, Biogen, Genzyme, Merck, Novartis, Sanofi, and UCB.

Multimedia Appendix 1

Trend estimation and residual modelling.

[PDF File (Adobe PDF File), 1MB - jmir_v15i1e20_app1.pdf]

References

1. de Lau LM, Breteler MM. Epidemiology of Parkinson's disease. The Lancet Neurology 2006 Jun;5(6):525-535. [doi: 10.1016/S1474-4422(06)70471-9] [Medline: 16713924]
2. Wirdefeldt K, Adami HO, Cole P, Trichopoulos D, Mandel J. Epidemiology and etiology of Parkinson's disease: a review of the evidence. Eur J Epidemiol 2011 Jun;26 Suppl 1:S1-58. [doi: 10.1007/s10654-011-9581-6] [Medline: 21626386]
3. Vu TC, Nutt JG, Holford NH. Progression of motor and nonmotor features of Parkinson's disease and their response to treatment. Br J Clin Pharmacol 2012 Aug;74(2):267-283. [doi: 10.1111/j.1365-2125.2012.04192.x] [Medline: 22283961]
4. Hoehn MM, Yahr MD. Parkinsonism: onset, progression and mortality. Neurology 1967 May;17(5):427-442. [Medline: 6067254]
5. Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease. The Unified Parkinson's Disease Rating Scale (UPDRS): status and recommendations. Mov Disord 2003 Jul;18(7):738-750 [FREE Full text] [doi: 10.1002/mds.10473] [Medline: 12815652]
6. Ramaker C, Marinus J, Stiggelbout AM, Van Hilten BJ. Systematic evaluation of rating scales for impairment and disability in Parkinson's disease. Mov Disord 2002 Sep;17(5):867-876. [doi: 10.1002/mds.10248] [Medline: 12360535]
7. Tsanas A, Little MA, McSharry PE, Scanlon BK, Papapetropoulos S. Statistical analysis and mapping of the Unified Parkinson's Disease Rating Scale to Hoehn and Yahr staging. Parkinsonism Relat Disord 2012 Jun;18(5):697-699. [doi: 10.1016/j.parkreldis.2012.01.011] [Medline: 22321863]
8. Martínez-Martín P, Gil-Nagel A, Gracia LM, Gómez JB, Martínez-Sarriés J, Bermejo F. Unified Parkinson's Disease Rating Scale characteristics and structure. The Cooperative Multicentric Group. Mov Disord 1994 Jan;9(1):76-83. [doi: 10.1002/mds.870090112] [Medline: 8139608]
9. PatientsLikeMe. URL: http://www.patientslikeme.com/ [accessed 2012-03-05] [WebCite Cache ID 65vSjbR4n]
10. Wicks P, Vaughan TE, Massagli MP, Heywood J. Accelerated clinical discovery using self-reported patient data collected online and a patient-matching algorithm. Nat Biotechnol 2011 May;29(5):411-414. [doi: 10.1038/nbt.1837] [Medline: 21516084]
11. Kurlan R, Murphy D. Parkinson's disease data and organizing center. Mov Disord 2007 Apr 30;22(6):904-905 [FREE Full text] [doi: 10.1002/mds.21415] [Medline: 17343272]
12. PatientsLikeMe. 2012 Mar 05. Privacy policy URL: http://www.patientslikeme.com/about/privacy [accessed 2012-03-07] [WebCite Cache ID 65zQdX8tJ]
13. Louis ED, Lynch T, Marder K, Fahn S. Reliability of patient completion of the historical section of the Unified Parkinson's Disease Rating Scale. Mov Disord 1996 Mar;11(2):185-192. [doi: 10.1002/mds.870110212] [Medline: 8684390]
14. Wasserman L. All of Statistics: A Concise Course in Statistical Inference. New York: Springer; 2004.
15. Davidian M, Giltinan DM. Nonlinear models for repeated measurement data: an overview and update. Journal of Agricultural, Biological, and Environmental Statistics 2003;8(4):387-419. [doi: 10.1198/1085711032697]
16. Bonate PL. Recommended reading in population pharmacokinetic pharmacodynamics. AAPS J 2005;7(2):E363-E373 [FREE Full text] [doi: 10.1208/aapsj070237] [Medline: 16353916]
17. Rice JA, Wu CO. Nonparametric mixed effects models for unequally sampled noisy curves. Biometrics 2001 Mar;57(1):253-259. [Medline: 11252607]
18. Muller P, Rosner GL, De Iorio M, MacEachern S. A nonparametric Bayesian model for inference in related longitudinal studies. Journal of the Royal Statistical Society: Series C (Applied Statistics) 2005;54(3):611-626. [doi: 10.1111/j.1467-9876.2005.05475.x]
19. Boyd S, Vandenberghe L. Convex Optimization. Cambridge, UK: Cambridge University Press; 2004.
20. Goetz CG, Blasucci L, Stebbins GT. Switching dopamine agonists in advanced Parkinson's disease: is rapid titration preferable to slow? Neurology 1999 Apr 12;52(6):1227-1229. [Medline: 10214748]
21. Wicks P, Massagli M, Kulkarni A, Dastani H. Use of an online community to develop patient-reported outcome instruments: the Multiple Sclerosis Treatment Adherence Questionnaire (MS-TAQ). J Med Internet Res 2011;13(1):e12 [FREE Full text] [doi: 10.2196/jmir.1687] [Medline: 21266318]
22. Shulman LM, Gruber-Baldini AL, Anderson KE, Fishman PS, Reich SG, Weiner WJ. The clinically important difference on the unified Parkinson's disease rating scale. Arch Neurol 2010 Jan;67(1):64-70. [doi: 10.1001/archneurol.2009.295] [Medline: 20065131]
23. Weiner WJ. Motor fluctuations in Parkinson's disease. Rev Neurol Dis 2006;3(3):101-108. [Medline: 17047575]
24. 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]
25. van Rooden SM, Colas F, Martínez-Martín P, Visser M, Verbaan D, Marinus J, et al. Clinical subtypes of Parkinson's disease. Mov Disord 2011 Jan;26(1):51-58. [doi: 10.1002/mds.23346] [Medline: 21322019]
26. Clifford GD, Clifton D. Wireless technology in disease management and medicine. Annu Rev Med 2012;63:479-492. [doi: 10.1146/annurev-med-051210-114650] [Medline: 22053737]
27. Tsanas A, Little MA, McSharry PE, Ramig LO. Nonlinear speech analysis algorithms mapped to a standard metric achieve clinically useful quantification of average Parkinson's disease symptom severity. J R Soc Interface 2011 Jun 6;8(59):842-855 [FREE Full text] [doi: 10.1098/rsif.2010.0456] [Medline: 21084338]

28. Kostikis N, Hristu-Varsakelis D, Arnaoutoglou M, Kotsavasiloglou C, Baloyiannis S. Towards remote evaluation of movement disorders via smartphones. Conf Proc IEEE Eng Med Biol Soc 2011;2011:5240-5243. [doi: 10.1109/IEMBS.2011.6091296] [Medline: 22255519]

Abbreviations

CID: clinically important difference
HY: Hoehn and Yahr
K-S: Kolmogorov-Smirnov test
PD: Parkinson's disease
PD-DOC: Parkinson's Disease Data and Organizing Center
PLM: PatientsLikeMe
q-q: quantile-quantile
UPDRS: Unified Parkinson's Disease Ratings Scale

Edited by G Eysenbach; submitted 16.03.12; peer-reviewed by D Kugiumtzis, K Dykstra; comments to author 12.06.12; revised version received 24.09.12; accepted 26.10.12; published 24.01.13.

Please cite as:
Little M, Wicks P, Vaughan T, Pentland A
Quantifying Short-Term Dynamics of Parkinson's Disease Using Self-Reported Symptom Data From an Internet Social Network
J Med Internet Res 2013;15(1):e20
URL: http://www.jmir.org/2013/1/e20/
doi:10.2196/jmir.2112
PMID:23343503

Original Paper

Correlates of Health-Related Social Media Use Among Adults

Rosemary Thackeray[1*], MPH, PhD; Benjamin T Crookston[1*], MPH, PhD; Joshua H West[1*], MPH, PhD

Brigham Young University, Department of Health Science, Provo, UT, United States

[*] all authors contributed equally

Corresponding Author:
Rosemary Thackeray, MPH, PhD
Brigham Young University
Department of Health Science
221 Richards Building
Provo, UT, 84602
United States
Phone: 1 801 422 1704
Fax: 1 801 422 0273
Email: rosemary_thackeray@byu.edu

Abstract

Background: Sixty percent of Internet users report using the Internet to look for health information. Social media sites are emerging as a potential source for online health information. However, little is known about how people use social media for such purposes.

Objectives: The purpose of this study was two-fold: (1) to establish the frequency of various types of online health-seeking behaviors, and (2) to identify correlates of 2 health-related online activities, social networking sites (SNS) for health-related activities and consulting online user-generated content for answers about health care providers, health facilities, or medical treatment.

Methods: The study consisted of a telephone survey of 1745 adults who reported going online to look for health-related information. Four subscales were created to measure use of online resources for (1) using SNS for health-related activities; (2) consulting online rankings and reviews of doctors, hospitals or medical facilities, and drugs or medical treatments; (3) posting a review online of doctors, hospitals or medical facilities, and drugs or medical treatments, and (4) posting a comment or question about health or medical issues on various social media. Univariate and multivariate logistic regression analyses were performed.

Results: Respondents consulted online rankings or reviews (41.15%), used SNS for health (31.58%), posted reviews (9.91%), and posted a comment, question, or information (15.19%). Respondents with a chronic disease were nearly twice as likely to consult online rankings (odds ratio [OR] 2.09, 95% CI 1.66-2.63, $P<.001$). Lower odds of consulting online reviews were associated with less formal education (OR 0.49, 95% CI 0.37-0.65, $P<.001$) and being male (OR 0.71, 95% CI 0.57-0.87, $P<.001$). Respondents with higher incomes were 1.5 times as likely to consult online rankings or reviews (OR 1.49, 95% CI 0.10-2.24, $P=.05$), than respondents with a regular provider (OR 2.05, 95% CI 1.52-2.78, $P<.001$), or living in an urban/suburban location (OR 1.61, 95% CI 1.17-2.22, $P<.001$). Older respondents were less likely to use SNS for health-related activities (OR 0.96, 95% CI 0.95-0.97, $P<.001$), as were males (OR 0.70, 95% CI 0.56-0.87, $P<.001$), whereas respondents with a regular provider had nearly twice the likelihood of using SNS for health-related activities (OR 1.89, 95% CI 1.43-2.52, $P<.001$).

Conclusions: People are using social media for seeking health information. However, individuals are more likely to consume information than they are to contribute to the dialog. The inherent value of "social" in social media is not being captured with online health information seeking. People with a regular health care provider, chronic disease, and those in younger age groups are more likely to consult online rankings and reviews and use SNS for health-related activities.

(J Med Internet Res 2013;15(1):e21) doi:10.2196/jmir.2297

KEYWORDS

social media; Internet; health information; consumer

Introduction

The Internet is becoming an increasingly common source of health information. Approximately 60% of Internet users report using the Internet to look for health information [1,2]. In addition to seeking health information, Wen et al [3] found that 15% of Internet users also tracked personal health information on the Internet. Determinants of seeking health information online include education, gender, race, age, presence of children in the home, having a poor personal health condition, and geographic residence [1,4-7]. Similarly, predictors of using the Internet to track personal health information include gender, race, education, and having a health care provider [3].

Historically, online health seeking meant visiting an agency- or organization-sponsored website. Recently, social media sites are emerging as a potential source of online health information [8]. Social media refers to “activities, practices, and behaviors among communities of people who gather online to share information, knowledge, and opinions using conversational media” [9]. These social media are broadly categorized as forums and message boards, review and opinion sites, social networks (eg, Facebook), blogging and microblogging (eg, Twitter), bookmarking, and media sharing (eg, YouTube) [10].

Individual use of social media is steadily increasing. Nearly two-thirds (65%) of adult Internet users in the United States are involved with a type of social media called *social networking sites* (SNS), such as MySpace, Facebook, or LinkedIn [11]. Technorati currently registers over 1.3 million blogs [12], 13 percent of Internet users (140 million people) have a Twitter account [13,14], and Facebook has 955 million active users [15]. By 2015, it is estimated that the number of individuals and corporations who have social networking accounts will reach over 3 billion [16].

Social media and SNS use varies by demographics. There are statistically significant differences in SNS use between younger and older ages and between males and females [11]. However, SNS are used fairly equally across education, income, race/ethnicity, and rural and urban locations [11]. Chou and colleagues [17] found that age and education were predictors of 3 forms of social media use (ie, participating in online support groups, blogging, and visiting a SNS).

In contrast to going online to seek health information, social media technologies allow online social media users to create, distribute, and share information independent of an organization. The level of use and involvement with social media technologies varies by individual. Bernoff and Anderson [18] and Li and Bernoff [19] classify individuals based on how they use social media. These classifications, although not mutually exclusive, include creators, conversationalists, critics, collectors, joiners, or spectators. Similarly, Hoffman and Novak [20] identify 4 goals for social media use: create, connect, consume, and control. The main conclusion from both typologies is that the range of social media activities that people engage in varies from consuming to creating content.

Despite the near ubiquity of social media use and the high prevalence of health information seeking on the Internet, there is a dearth of literature about the characteristics of people who use social media for seeking health information and how these people engage with social media. Thus, additional research is needed to determine whether social media users are primarily spectators, or if they are creators or critics. That is, are they looking for information or are they becoming part of the information creation and sharing process? Knowing the correlates of social media use for health information can allow health professionals to more accurately segment populations and tailor interventions accordingly. Therefore, the aim of this research was two-fold. First, to establish the frequency of various forms (eg, spectators, creators, or critics) of online health-seeking behaviors. Second, this research seeks to identify correlates of 2 health-related online activities: (1) using SNS for health-related activities, and (2) consulting online user-generated content for answers about health care providers, health facilities, or medical treatment.

Methods

Data Source and Sample

The data for this study were taken from the 2010 Health Tracking Survey conducted by Princeton Survey Research on behalf of the Pew Internet & American Life Project [21]. The data were collected during August to September 2010 through a telephone survey that included both cell phones and landlines. A random digit method was used to select participants who were US residents, aged 18 years and older, and who spoke English (n=3001). Data were weighted to the most recent US Census Bureau’s Current Population Survey. Data were stripped of identifying information and made available to the public. For the current study, inclusion criteria were adults who used the Internet at least occasionally (Pew question Q6a) and who reported going online to look for health-related information (Pew question healthseek). The final sample size was 1745.

Measures

Demographic, Socioeconomic, and Health Status

Demographic, socioeconomic, and health status covariates included ethnicity, education, income, gender, age, race, marital status, having a chronic health condition, geographic community type, health insurance status, and having a family doctor or health care professional. Response categories for race, education, and marital status were collapsed to account for small cell sizes.

To measure an individual’s level of social media health engagement, we created 4 subscales based on related survey items. Each of the response variables were dichotomous and coded as yes or no. We calculated Cronbach alpha to estimate internal reliability for each scale.

Used Social Networking Sites for Health-Related Activities

The 5 questions that focused on using SNS for health-related activities included (1) get health information, (2) start or join a health-related group, (3) follow your friend’s personal health experiences or health updates, (4) raise money or draw attention to a health-related issue or cause, and (5) remember or memorialize others who suffered from a certain health condition

(Pew questions Q26a-e). The composite scale had an internal reliability of Cronbach alpha=.66

Consulted Online Rankings or Reviews

Three questions focused on consulting online rankings or reviews of (1) doctors or other providers, (2) hospitals or other medical facilities, and (3) particular drugs or medical treatments (Pew questions Q29a-c; Cronbach alpha=.69).

Posted a Review Online

Three questions focused on whether respondents had posted a review online of (1) a doctor, (2) a hospital, or (3) his/her experiences with a particular drug or medical treatment (Pew questions Q29d-f; Cronbach alpha=.61)

Posted a Comment or Question on Social Media

Five questions asked if respondents had posted comments, questions, or information about health or medical issues on various social media. These included (1) an online discussion, a listserv, or other online group forum, (2) a blog, (3) a social networking site, such as Facebook, MySpace, or LinkedIn, (4) Twitter or another status update site, and (5) a website of any kind, such as a health site or news site that allows comments and discussion (Pew questions Q25a-e; Cronbach alpha=.80)

Data Analysis

Unadjusted univariate analyses of demographics, socioeconomic, and health status variables with each social media–health engagement scale were computed. Variables that were significantly associated with the dependent variable were included in a multivariate regression model. Multivariate logistic regression analysis was performed with social media health engagement as the dependent variable and the demographics, socioeconomic, and health status variables as covariates. All analyses were conducted using IBM SPSS Statistics version 20 (IBM Corp, Armonk, NY, USA).

Results

Demographic Characteristics

More than half of the study sample was female (56.16%, 980/1745) and white (79.20%, 1382/1745) (see Table 1). College graduates comprised 39.43% (689/1745) of the sample, 29.46% (514/1745) reported a household income between US $75,000 and $150,000, and 86.88% (1516/1745) reported having health insurance. Respondents reported consulting online rankings or reviews (41.15%, 718/1745) and using SNS for health (31.58%, 551/1745) more than they reported contributing content through posting reviews of doctors, hospitals, drugs, or medical treatments (9.91%, 173/1745), or posting a comment, question, or information about health or medical issues on a blog, SNS, Twitter, website, or online discussion or forum (15.19%, 265/1745).

Correlates of Social Media Health Engagement and Regression Analyses

Regression analyses revealed few correlates for posting reviews of a doctor, hospital, drug, or medical treatment (chronic disease, income, age, health insurance) and for posting a comment, question, or information on various social media sites (chronic disease, age, marital status). Therefore, the further analysis and data presented here are limited to using SNS for health and consulting online rankings or reviews.

An examination of correlates of consulting online rankings or reviews identifies several factors that are associated with higher use of online rankings and reviews (see Table 2). For example, approximately half (49.27%, 339/688) of those with a college degree reported using online rankings or reviews compared with 40.71% (204/501) of those with some college and 31.50% (172/546) of those with a high school education or less. Factors associated with use of SNS for health included income, gender, age, marital status, and having a personal or family doctor or health care provider.

Unadjusted Regression Analyses for Consulting Online Rankings

Unadjusted regression analyses revealed numerous factors associated with consulting online rankings or reviews of doctors, hospitals, drugs, or medical treatments (see Table 3). Having a chronic disease, reporting a higher annual income, living in an urban/suburban location, reporting health insurance coverage, and having a regular health care provider were each independently associated with increased odds of consulting online rankings. Decreased odds were observed among older respondents, those who were unmarried, those with lower levels of education, males, and those who were black/African American.

Unadjusted Regression Analyses for Using Social Networking Sites for Health

Table 4 presents the results of the unadjusted regression analyses for using SNS for health-related activities, such as getting information, joining a group, following friends' health experiences, raising money, increasing awareness, or remembering or memorializing others. Older respondents and males were each less likely to engage in such behaviors. Respondents who reported being unmarried or having a regular health care provider were more likely to use SNS for health-related purposes.

Adjusted Regression Analyses for Consulting Online Rankings

Results from adjusted regression analyses (see Table 3) revealed that respondents with a chronic disease were nearly twice as likely to consult online rankings as respondents who were free of chronic disease (OR 2.09, 95% CI 1.66-2.63, *P*<.001). For levels of education, high school or less (OR 0.49, 95% CI 0.37-0.66, *P*<.001) and some college (OR 0.70, 95% CI 0.54-0.91, *P*=.01) were each associated with lower odds of consulting online rankings than respondents who had at least obtained a college degree. With respect to income, respondents who reported an annual income of US $75,000 to $150,000 were 1.5 times as likely to consult online rankings (OR 1.49, 95% CI 0.10-2.24, *P*=.05) compared to those making less than US $20,000. Males were less likely than females (OR 0.71, 95% CI 0.57-0.87, *P*<.001), whereas respondents who have a regular provider were more than 2 times more likely to consult online rankings (OR 2.05, 95% CI 1.52-2.78, *P*<.001). Living in an urban/suburban location was associated with a 60%

increased chance of consulting rankings (OR 1.61, 95% CI 1.17-2.22, *P*<.001). In the adjusted model, marital status, race, and insurance coverage were not significantly associated with consulting online rankings. Likewise, the influence of having a health care provider and income was attenuated.

Adjusted Regression Analyses for Using Social Networking Sites for Health

Adjusted odds ratios for using SNS for health-related purposes are presented in Table 4. As respondents' ages increased, their likelihood for using such sites decreased (OR 0.96, 95% CI 0.95-0.97, *P*<.001). With respect to gender, males had lower odds than females (OR 0.70, 95% CI 0.56-0.87, *P*<.001). Compared to respondents without a regular health care provider, respondents with a regular provider had nearly twice the likelihood of using SNS for health-related activities (OR 1.89, 95% CI 1.43-2.52, *P*<.001), a greater influence than in the unadjusted model. Marital status was not significantly associated with using SNS for health-related activities.

Table 1. Sample characteristics (N=1745).

Characteristic	n (%)
Gender	
Male	765 (43.84)
Female	980 (56.16)
Education	
Less than high school or high school graduate	546 (31.29)
Some college	501 (28.71)
College graduate or more	688 (39.43)
Answer not give	10 (0.57)
Income (US$)	
<$20,000	228 (13.07)
$20,000 to <$40,000	347 (19.88)
$40,000 to <$75,000	429 (24.58)
$75,000 to <$150,000	514 (29.46)
Don’t know or refused	227 (13.01)
Race	
White	1382 (79.20)
Black/African American	195 (11.17)
All other races	136 (7.79)
Answer not given	32 (1.83)
Marital status	
Married or living with a partner	1060 (60.74)
All other marital status	676 (38.74)
Answer not given	9 (0.52)
Geographic community type	
Rural	228 (13.07)
Urban/suburban	1451 (83.15)
Answer not given	66 (3.78)
Has a chronic health condition	
Yes	699 (40.06)
No	1046 (59.94)
Has a regular health care provider	
Yes	1398 (80.11)
No	343 (19.66)
Answer not given	4 (0.23)
Health insurance	
Yes	1516 (86.88)
No	229 (13.12)
Hispanic ethnicity	
Yes	182 (10.43)
No	1552 (88.94)
Answer not given	11 (0.63)
Post health-related comments or questions on 5 social media	

Characteristic	n (%)
Yes	265 (15.19)
No	1480 (84.81)
Use social networking sites for health-related information	
Yes	551 (31.58)
No	1194 (68.42)
Consulted online rankings or reviews of doctors, hospitals, drugs, or medical treatment	
Yes	718 (41.15)
No	1027 (58.85)
Posted a review of doctors, hospitals, drugs, or medical treatment	
Yes	173 (9.91)
No	1572 (90.09)

[a] The 5 social media are online discussion, listserv or other online group forum, a blog, a social networking site, Twitter or another status update site, and a website of any kind.

Table 2. Correlates of consulting online rankings or reviews (n=718) and use of social networking sites for health (n=551).

Sociodemographic and health characteristics		Total sample[a] (N=1745)	Consulted online rankings or reviews		Use social networking sites for health	
			n (%)	*P*	n (%)	*P*
Ethnicity				.20		.26
	Hispanic/Latino	183	67 (36.61)		51 (27.87)	
	Non-Hispanic/Latino	1552	644 (41.49)		499 (32.15)	
Education				<.001		.21
	Less than high school diploma or a high school graduate	546	172 (31.50)		160 (29.30)	
	Some college	501	204 (40.71)		156 (31.14)	
	College degree or more	689	339 (49.37)		234 (33.96)	
Chronic health condition				<.001		.13
	Yes	699	347 (49.64)		206 (29.47)	
	No	1046	371 (35.47)		345 (32.95)	
Income (US$)				<.001		.03
	<$20,000	228	76 (33.33)		65 (28.51)	
	$20,000 to <$40,000	347	131 (37.75)		119 (34.29)	
	$40,000 to <$75,000	429	167 (38.93)		149 (34.73)	
	$75,000 to <$150,000	514	255 (49.61)		165 (32.10)	
	Don’t know/refused	227	88 (38.77)		54 (23.79)	
Gender				.001		.001
	Male	765	280 (36.60)		215 (28.10)	
	Female	979	437 (44.64)		335 (34.22)	
Age				<.001		<.001
Geographic community type				.001		.09
	Rural	229	75 (32.75)		61 (26.64)	
	Urban/suburban	1451	616 (42.45)		468 (32.25)	
Health insurance				.04		.64
	Yes	1516	654 (43.14)		475 (31.33)	
	No	229	64 (27.95)		75 (32.75)	
Race				.04		.52
	White	1382	588 (42.55)		444 (32.13)	
	Black/African American	195	68 (34.87)		63 (32.31)	
	All other races	136	47 (34.56)		37 (27.21)	
Marital status				.001		< .001
	Married or living with a partner	1060	468 (44.15)		284 (26.79)	
	All other marital status	676	244 (36.09)		264 (39.05)	
Personal or family doctor or health care professional				<.001		.047
	Yes	1398	632 (45.21)		457 (32.69)	
	No	343	83 (24.20)		93 (27.11)	

[a] Total n may not equal 1745 due to missing values and rounding of weighted data.

Table 3. Consulted online rankings or reviews of doctors, hospitals, drugs, or medical treatments.

Sociodemographic and health characteristics	Unadjusted regression			Adjusted regression		
	Odds ratio (SE)	95% CI	*P*	Odds ratio (SE)	95% CI	*P*
Chronic disease						
Does not have a chronic disease	1.0			1.0		
Has a chronic disease	1.79 (0.10)	1.48-2.18	<.001	2.09 (0.10)	1.66-2.63	<.001
Age	0.96 (0.00)		<.001	0.99 (0.00)	0.98-0.10	.01
Marital status						
Married	1.0			1.0		
All other marital status	0.72 (0.10)	0.59-0.87	<.001	0.87 (0.13)	0.68-1.10	.26
Education						
Less than high school graduate or high school graduate	0.48 (0.12)	0.38-0.60	<.001	0.49 (0.14)	0.37-0.65	<.001
Some college	0.71 (0.012)	0.56-0.89	<.001	0.70 (0.13)	0.54-0.91	.01
College degree or higher	1.0			1.0		
Income (US$)						
<$20,000	1.0			1.0		
$20,000 to <$40,000	1.21 (0.18)	0.85-1.72	0.28	1.20 (0.20)	0.82-1.79	.15
$40,000 to <$75,000	1.28 (0.17)	0.91-1.79	.15	1.07 (0.20)	0.73-1.58	.72
$75,000 to <$150,000	1.98 (0.17)	1.43-2.73	<.001	1.49 (0.21)	0.10-2.24	.05
Don’t know/refused	1.27 (0.20)	0.87-1.87	.21	1.09 (0.23)	0.70-1.71	.70
Gender						
Female	1.0			1.0		
Male	0.72 (0.10)	0.59-0.87	<.001	0.71 (0.11)	0.57-0.87	<.001
Geographic location						
Urban/suburban	1.59 (0.15)	1.13-2.04	.01	1.61 (0.16)	1.17-2.22	<.001
Rural	1.0			1.0		
Race						
White	1.0			1.0		
Black/African American	0.72	0.52-0.98	.04	0.79 (0.18)	0.55-1.12	.19
All other races	0.70	0.49-1.02	.06	0.69 (0.21)	0.46-1.03	.075
Insurance coverage						
Does not have health insurance	1.0			1.0		
Has health insurance	1.95 (0.16)	1.44-2.65	<.001	1.16 (0.18)	0.81-1.65	.41
Regular health care provider						
Does not have a regular health care provider	1.0			1.0		
Has a regular health care provider	2.59 (0.14)	1.98-3.39	<.001	2.05 (0.15)	1.52-2.78	<.001

Table 4. Used social networking sites for health-related activities, such as getting information, joining a group, following friends' health experiences, raising money, increasing awareness, and remembering or memorializing others.

Sociodemographic and health characteristics	Unadjusted			Adjusted		
	Odds ratio (SE)	95% CI	*P*	Odds ratio (SE)	95% CI	*P*
Age	0.96 (0.00)	0.96-0.97	<.001	0.96 (0.00)	0.95-0.97	<.001
Marital status						
Married	1.0			1.0		
All other marital statuses	1.45 (0.11)	1.42-2.15	<.001	1.18 (0.12)	0.94-1.49	.16
Gender						
Female	1.0			1.0		
Male	0.75 (0.11)	0.61-0.93	.01	0.70 (0.11)	0.56-0.87	<.001
Regular health care provider						
Does not have a regular health care provider	1.0			1.0		
Has a regular health care provider	1.31 (0.13)	1.00-1.70	.048	1.89 (0.15)	1.43-2.52	<.001

Discussion

This study examined the frequency of engaging in content creation through posting on social media sites, consumption of online rankings and reviews, and use of SNS for health-related activities. In addition, correlates for engaging in those behaviors were examined. The rate of online health information seeking behavior was similar to what has been reported previously [1,2]. Results show that although social media technologies allow people the opportunity to participate in the creation of online information, this is not very common; less than 15% of people reported doing so. In contrast, people are more likely to consume content with 30% to 40% of respondents reporting use of SNS for health-related activities and use of online rankings or review of doctors, hospitals, and medical treatments.

The lack of creating and contributing content is an intriguing finding. The value of social media is in the sharing of information within social networks. The rate of contributing opinions and experiences on other social media venues, such as product reviews or rankings sites (eg, Amazon or TripAdvisor), is similar to what we found in the current study [22]. People are not contributors. One explanation might be related to the fact that the frequency of encounters with doctors, hospitals, or medical treatments is less often, so there may be less motivation to share experiences. It could also be due to users' feelings of incompetence relating to health topics, preferring to leave such discussions to trained professionals. Overall, there is a need for more research to understand the motivations and perceived benefits of contributing to health-related online forums, discussion boards, rating sites, and other social media venues.

Use of SNS for health was more common among females and younger people. These findings are not surprising given that this same group is more likely to use SNS in general [11]. Although growing in popularity among older populations, SNS use is still more common among people younger than 50 years of age, and particularly among the 18-29 year age group [11], which is consistent with our findings that younger audiences are more likely to use SNS for health-related activities.

People with chronic disease were twice as likely to consult online rankings or reviews. Previous research has shown that people in poorer health are more likely to seek health information online [6,23], hence consulting online data about doctors, hospitals, and medical treatments is probably reflective of the need for information to manage their condition. In addition, people who more frequently use health care services may be more invested in their health and, therefore, seek high quality experiences. Additionally, individuals with a chronic disease may also have greater medical knowledge about their condition and may feel more competent sharing that knowledge in a social media venue. It may be that use of social media for health is most applicable for specific segments of the population, such as those people who are trying to manage a chronic health condition.

Higher income was also associated with increased likelihood of consulting online rankings and reviews. This is similar to research that found health information seeking was also more common among higher income groups [24]. This may be because people with higher income use the Internet more often [25]. Alternatively, more health care options may be available to those with higher income than to those with lower income because of health insurance coverage, so they are able to discriminate among their choices in providers and treatment.

Regression models showed that having a regular health care provider is the only significant variable associated with both consulting online rankings and reviews and for using SNS for health-related activities. This is consistent with previous research that found that having a health care provider is associated with tracking personal health information on the Internet [26]. It might also be an artifact of higher income; individuals with a regular health care provider may also be wealthier. But the findings are inconsistent with Chou et al [17] who found that having a health care provider was not associated with social media use for health-related purposes, primarily due to age.

Variables that are traditionally associated with online health-seeking behavior, including race, geography, health insurance coverage, marital status, and education were not significant in terms of consulting online rankings for reviews or using SNS for health, both a type of health-seeking behavior. Thus, using social media for health seeking may be less influenced by common sociodemographic variables and may be better explained by other factors. For example, research on social network use has shown that personality traits, such as extroversion and neuroticism, are associated with social media use and sharing of information [27,28].

Limitations

The data should be interpreted with caution considering the following data limitations. The internal consistency for the scales use to measure social media use ranged from .609 to .798. Although these are acceptable values according to conventional research standards, the individual items may not accurately capture the array of health behaviors one may engage in while using social media. For instance, people may be posting reviews online about other health-related experiences than the 3 assessed by this survey. Adding more variables would increase the internal consistency. These variables would have to be added to the Pew survey and may include items such as posting or consulting reviews of community or nonprofit facilities where services were received, using social networking sites to track personal progress toward health-related goals, or to receive social support, and so forth.

Conclusions

People are using social media for seeking health information. However, individuals are more likely to consume information than they are to contribute to the dialog. The inherent value of “social” in social media is not being captured with online health information seeking. People with a regular health care provider, chronic disease, and those in younger age groups are more likely to consult online rankings and reviews and use SNS for health-related activities.

Acknowledgments

The Pew Research Center for the People & the Press is the source of the data. The Center bears no responsibility for the interpretations presented or conclusions reached based on analysis of the data.

Conflicts of Interest

None declared.

References

1. Atkinson NL, Saperstein SL, Pleis J. Using the internet for health-related activities: findings from a national probability sample. J Med Internet Res 2009;11(1):e4 [FREE Full text] [doi: 10.2196/jmir.1035] [Medline: 19275980]
2. Ybarra M, Suman M. Reasons, assessments and actions taken: sex and age differences in uses of Internet health information. Health Educ Res 2008 Jun;23(3):512-521 [FREE Full text] [doi: 10.1093/her/cyl062] [Medline: 16880222]
3. Wen KY, Kreps G, Zhu F, Miller S. Consumers' perceptions about and use of the internet for personal health records and health information exchange: analysis of the 2007 Health Information National Trends Survey. J Med Internet Res 2010;12(4):e73 [FREE Full text] [doi: 10.2196/jmir.1668] [Medline: 21169163]
4. Lustria ML, Smith SA, Hinnant CC. Exploring digital divides: an examination of eHealth technology use in health information seeking, communication and personal health information management in the USA. Health Informatics J 2011 Sep;17(3):224-243. [doi: 10.1177/1460458211414843] [Medline: 21937464]
5. Ybarra ML, Suman M. Help seeking behavior and the Internet: a national survey. Int J Med Inform 2006 Jan;75(1):29-41. [doi: 10.1016/j.ijmedinf.2005.07.029] [Medline: 16129659]
6. Rice RE. Influences, usage, and outcomes of Internet health information searching: multivariate results from the Pew surveys. Int J Med Inform 2006 Jan;75(1):8-28. [doi: 10.1016/j.ijmedinf.2005.07.032] [Medline: 16125453]
7. Ruggiero KJ, Gros DF, McCauley J, de Arellano MA, Danielson CK. Rural adults' use of health-related information online: data from a 2006 National Online Health Survey. Telemed J E Health 2011 Jun;17(5):329-334 [FREE Full text] [doi: 10.1089/tmj.2010.0195] [Medline: 21524201]
8. Fox S. The social life of health information, 2011. Washington, DC: Pew Internet & American Life Project; 2011 May 12. URL: http://pewinternet.org/reports/2011/social-life-of-health-info.aspx [accessed 2012-08-01] [WebCite Cache ID 69bCrjKnN]
9. Safko J, Brake DK. The Social Media Bible: Tactics, Tools, and Strategies for Business Success. Hoboken, NJ: Wiley; 2009.
10. Stern J. Social Media Metrics: How to Measure and Optimize Your Marketing Investment. Hoboken, NJ: John Wiley; 2010.
11. Madden M, Zickuhr K. 65% of online adults use social networking sites. Washington, DC: Pew Internet & American Life Project; 2011 Aug 26. URL: http://www.pewinternet.org/Reports/2011/Social-Networking-Sites.aspx [accessed 2012-08-01] [WebCite Cache ID 69bCvF51S]
12. Technorati blog directory. 2011. URL: http://technorati.com/blogs/directory/ [accessed 2012-08-01] [WebCite Cache ID 69bCmAOWC]

13. Smith A. 13% of online adults use Twitter. Washington, DC: Pew Internet & American Life Project; 2011 Jun 01. URL: http://www.pewinternet.org/~/media//Files/Reports/2011/Twitter%20Update%202011.pdf; [accessed 2012-08-01] [WebCite Cache ID 69bCxrzY9]
14. Twitter. Twitter turns six. 2012 Mar 26. URL: http://blog.twitter.com/2012/03/twitter-turns-six.html [accessed 2012-08-01] [WebCite Cache ID 69bD02ViQ]
15. Facebook. Key facts URL: http://newsroom.fb.com/content/default.aspx?NewsAreaId=22 [accessed 2012-08-01] [WebCite Cache ID 69bCiROYY]
16. The Radicati Group, Inc. Email statistics report, 2011-2015. 2011 May 18. URL: http://www.radicati.com/?p=7261 [accessed 2012-08-01] [WebCite Cache ID 69bD3bbJc]
17. Chou WY, Hunt YM, Beckjord EB, Moser RP, Hesse BW. Social media use in the United States: implications for health communication. J Med Internet Res 2009;11(4):e48 [FREE Full text] [doi: 10.2196/jmir.1249] [Medline: 19945947]
18. Bernoff J, Anderson J. Social techographics defined. 2010. URL: http://www.forrester.com/error/?eCode=490 [accessed 2012-08-01] [WebCite Cache ID 69bD5Zwiu]
19. Li C, Bernoff J. Groundswell: Winning in a World Transformed by Social Technologies. Boston, MA: Harvard Business Press; 2008.
20. Hoffman DL, Novak TP. Social Science Research Network. 2012. Why do people use social media? Empirical findings and a new theoretical framework for social media goal pursuit URL: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1989586 [accessed 2012-08-01] [WebCite Cache ID 69bjzTpnc]
21. Pew Internet. Washington, DC: Pew Internet & American Life Project; 2012. Our research URL: http://pewinternet.org/about-us/our-research [accessed 2012-08-01] [WebCite Cache ID 69bCWFezV]
22. Bronner F, de Hoog R. Vacationers and eWOM: Who posts, and why, where, and what? Journal of Travel Research 2011;50(1):15-26. [doi: 10.1177/0047287509355324]
23. Choi N. Relationship between health service use and health information technology use among older adults: analysis of the US National Health Interview Survey. J Med Internet Res 2011;13(2):e33 [FREE Full text] [doi: 10.2196/jmir.1753] [Medline: 21752784]
24. Beaudoin CE, Hong T. Health information seeking, diet and physical activity: an empirical assessment by medium and critical demographics. Int J Med Inform 2011 Aug;80(8):586-595. [doi: 10.1016/j.ijmedinf.2011.04.003] [Medline: 21640643]
25. Jansen J. Use of the Internet in higher-income households. Washington, DC: Pew Internet & American Life Project; 2010 Nov 24. URL: http://www.pewinternet.org/Reports/2010/Better-off-households.aspx [accessed 2012-08-01] [WebCite Cache ID 69bD7x4PS]
26. Wen LM, Rissel C, Baur LA, Lee E, Simpson JM. Who is NOT likely to access the Internet for health information? Findings from first-time mothers in southwest Sydney, Australia. Int J Med Inform 2011 Jun;80(6):406-411. [doi: 10.1016/j.ijmedinf.2011.03.001] [Medline: 21493126]
27. Correa T, Hinsley AW, de Zúñiga HG. Who interacts on the web?: The intersection of users' personality and social media use. Comput Hum Behav 2010;26(2):247-253. [doi: 10.1016/j.chb.2009.09.003]
28. Amichai-Hamburger Y, Vinitzky G. Social network use and personality. Comput Hum Behav 2010;26(6):1289-1295. [doi: 10.1016/j.chb.2010.03.018]

Abbreviations

SNS: social networking sites

Edited by G Eysenbach; submitted 02.08.12; peer-reviewed by T Hartz; comments to author 06.10.12; revised version received 11.10.12; accepted 17.10.12; published 30.01.13.

Please cite as:
Thackeray R, Crookston BT, West JH
Correlates of Health-Related Social Media Use Among Adults
J Med Internet Res 2013;15(1):e21
URL: http://www.jmir.org/2013/1/e21/
doi:10.2196/jmir.2297
PMID:23367505

Original Paper

Accessing Suicide-Related Information on the Internet: A Retrospective Observational Study of Search Behavior

Paul Wai-Ching Wong[1,2*], PsyD; King-Wa Fu[2,3*], Ph.D; Rickey Sai-Pong Yau[2*], BEng; Helen Hei-Man Ma[1,2*], MA; Yik-Wa Law[1,2*], Ph.D; Shu-Sen Chang[2*], MD, PhD; Paul Siu-Fai Yip[1,2*], PhD

[1]Department of Social Work and Social Administration, Faculty of Social Sciences, The University of Hong Kong, Pokfulam, China (Hong Kong)
[2]Centre for Suicide Research and Prevention, Faculty of Social Sciences, The University of Hong Kong, Hong Kong SAR, China (Hong Kong)
[3]Journalism and Media Studies Centre, The University of Hong Kong, Hong Kong SAR, China (Hong Kong)
[*]all authors contributed equally

Corresponding Author:
Paul Wai-Ching Wong, PsyD
Department of Social Work and Social Administration
Faculty of Social Sciences
The University of Hong Kong
Room 511, The Jockey Club Tower, The Centennial Campus,
The University of Hong Kong
Pokfulam,
China (Hong Kong)
Phone: 852 3917 5029
Fax: 852 2858 7604
Email: paulw@hku.hk

Abstract

Background: The Internet's potential impact on suicide is of major public health interest as easy online access to pro-suicide information or specific suicide methods may increase suicide risk among vulnerable Internet users. Little is known, however, about users' actual searching and browsing behaviors of online suicide-related information.

Objective: To investigate what webpages people actually clicked on after searching with suicide-related queries on a search engine and to examine what queries people used to get access to pro-suicide websites.

Methods: A retrospective observational study was done. We used a web search dataset released by America Online (AOL). The dataset was randomly sampled from all AOL subscribers' web queries between March and May 2006 and generated by 657,000 service subscribers.

Results: We found 5526 search queries (0.026%, 5526/21,000,000) that included the keyword "suicide". The 5526 search queries included 1586 different search terms and were generated by 1625 unique subscribers (0.25%, 1625/657,000). Of these queries, 61.38% (3392/5526) were followed by users clicking on a search result. Of these 3392 queries, 1344 (39.62%) webpages were clicked on by 930 unique users but only 1314 of those webpages were accessible during the study period. Each clicked-through webpage was classified into 11 categories. The categories of the most visited webpages were: entertainment (30.13%; 396/1314), scientific information (18.31%; 240/1314), and community resources (14.53%; 191/1314). Among the 1314 accessed webpages, we could identify only two pro-suicide websites. We found that the search terms used to access these sites included "commiting suicide with a gas oven", "hairless goat", "pictures of murder by strangulation", and "photo of a severe burn". A limitation of our study is that the database may be dated and confined to mainly English webpages.

Conclusions: Searching or browsing suicide-related or pro-suicide webpages was uncommon, although a small group of users did access websites that contain detailed suicide method information.

(J Med Internet Res 2013;15(1):e3) doi:10.2196/jmir.2181

KEYWORDS

Internet search; Pagerank; Suicide information; Information seeking; Search behavior; Information retrieval

Introduction

Internet usage is growing exponentially with one third of the world population having Internet access as of 2011, of which 45% were users below the age of 25. The number of Internet users from developing countries has increased from 44% in 2006 to 62% in 2011 [1]. The proliferation and reach of the Internet and its impact on suicide and self-harm behaviors has recently come to the attention of the general public, especially to parents of children and young individuals, suicide prevention professionals, and governments [2]. There is little evidence, however, to either support or refute the notion that the use of the Internet is intrinsically damaging [3].

Infodemiology is a new area of research described as "the science of distribution and determinants of information in an electronic medium," [4] specifically, the Internet, to inform public health and policy [5]. Recent applications of this methodology include "infoveillance," the tracking and surveillance of online data for the purpose of predicting influenza outbreaks [6] and the online spread of smoking cessation awareness [7], as well as investigating public search trends in preventing avian flu [8]. The real-time nature of this methodology means that results are timely and may have significant impact in guiding policy making.

In the area of suicide prevention, Dobson [9] has estimated that more than 100,000 suicide-related websites were available on the Internet but due to many search engines' conservative policies regarding the exposure or sharing of corporate data, there are no reliable data on the prevalence of websites with suicide-related content [10]. The major concerns about the potential role of Internet on suicidal behavior are: 1) the Internet as a major source of readily accessible information on suicide methods and suicidal communication, 2) access to suicide-related information and discussions online by vulnerable, impressionable individuals (eg, the distressed, stressed, and depressed) [11], and young people [12,13] contributes to the increased risk of suicide attempts and completion for them as ways to solve problems or to regulate negative emotions [14,15], and 3) the Internet as an efficient and communal medium for people to meet and arrange suicide pacts [16-18]. Due to the interactive nature of the Internet, Becker and Schmidt [19] argued that the Internet might have even more impact on copycat suicides than print media, eg, the Sorrows of Young Werther. This is the influential book from which the term "Werther effect" arose, which describes the effect of a widely publicized suicide causing emulated suicides.

Search engines on the Internet are often assumed to be an efficient, if not the preferred, gateway for access to suicide-related information [2,20]. A handful of infodemiological studies have been conducted to examine the types of information suicidal individuals might find on the Internet through search engines [15,21-23]. A similar "supply-based" research methodology [4] was adopted in these studies where various self-chosen suicide-related keywords (eg, "suicide", "suicide methods", "ways to commit suicide", or "how to kill yourself") were entered into several popular search engines, ie, Google, Yahoo!, and the search results were classified into pro-suicide, anti-suicide, or neutral websites. Around 11%, 25%, and 5% of the search result websites were categorized as pro-suicide in the United States, the United Kingdom, and China, respectively [21-23]. Based on these findings, the researchers advocated a number of potential suicide prevention initiatives: 1) Internet service providers and website hosting companies should do more to remove pro-suicide websites [23], 2) Internet service providers may consider strategies to increase the likelihood that vulnerable individuals access helpful rather than potentially harmful websites in times of crisis [21], and 3) enhance access to anti-suicide websites [20] in order to minimize the risks and maximize the benefits of the effects that search engines may have on suicide and suicide prevention. Indeed, a number of national-level policies or initiatives have been implemented in Australia (ie, making the promotion of suicide on the Internet illegal), and in Japan and Korea (ie, specific pro-suicide websites blocked by Internet service providers) [21].

Although prior research has shown that pro-suicide information on methods of suicide are easily accessible through search engines, the findings invited the criticism that the keywords and phrases used were pre-selected by researchers and thus had poor generalizability since the studies were not grounded in the naturalistic search behavior of non-researchers, web-savvy, or suicidal people [24]. Anecdotal evidence has been reported in the literature showing that people who have engaged in attempted suicides had searched for specific ways of committing suicide through the Internet, eg, [25]. There is no empirical data, however, on the general public's actual web searching behavior for suicide-related information through search engines. This information is needed to generate empirically informed suicide research and prevention strategies on the Internet. The current infodemiological study aims to address this research gap by using a demand-based method [4] investigating the actual searching behavior for suicide-related information on search engines using a publicly accessible dataset released by the America Online (AOL). We examined the search queries people used to access pro-suicide websites through AOL's search engine. Two research questions regarding suicide and the Internet were examined: 1) What websites do people access to when they search using queries containing "suicide"?, and 2) What search queries (both suicide and non-suicide included) do people use that result in access to pro-suicide websites?

Methods

Design

A web search dataset released by the AOL is publicly accessible for data analysis (see similar research in [26]). This dataset is currently the largest publicly accessible search query log on the Internet and represented approximately 1.5% of the total number of search queries conducted through the AOL search engine, which had a 6.7% share of the overall search market in 2006 [27]. The dataset was a random sampling from all AOL subscribers' search queries between March and May 2006 and consisted of roughly 21 million instances of such queries, collected from roughly 657,000 AOL subscribers [28].

Each entry contained an anonymous service subscriber identity code (ID), time of submission, the rank position, and domain of the clicked-through webpage. A single search query represents one individual click on the "search" button as initiated by a user, regardless of whether the user clicked on any of the webpages in the results page(s). A unique set of one or more keywords used in a search query constitutes the search term(s). A click-through is where a webpage in the returned search results was accessed by an individual click following a single instance of a search query and consequently that webpage is referred to as a clicked-through webpage. The number of click-throughs can also be referred to as traffic.

In response to concerns about privacy, additional measures were taken to replace all possible personal identifiable information from the dataset before undertaking analysis. Each unique subscriber was assigned a random identifier number. The Human Research Ethics Committee for Nonclinical Faculties of the University of Hong Kong approved the ethical aspect of the study protocol.

Data Coding

To address our first research question, any search query containing the keyword *suicide* in the search term(s) was retrieved from the dataset, and the contents of all clicked-through webpages following these search queries were then content analyzed to derive categories for analysis. Researchers PW and RY independently categorized all of the identified webpages, and each created an initial coding frame. The two coding frames were then compared and refined to develop a consensus coding frame. All of the clicked-through webpages were subsequently individually re-categorized by researchers PW and RY to form the final coding frame. Using the final coding frame, coding disagreements between PW and RY that could not be resolved were referred to the third researcher KWF, who analyzed those blinded to the codes assigned by researchers PW and RY. In most cases, this approach resolved the disagreements but when a decision still could not be made, researchers PW, RY, and KWF jointly visited the clicked-through webpage and discussed the rationale for their coding until agreement was reached.

To address our second research question, web domains from the webpages in the above 11 categories that appeared to encourage suicide or promote the individual right of dying by suicide regardless of medical grounds [23] were identified and content analyzed, and all the search terms used to access these pro-suicide web domains, including ones that did not contain the keyword suicide and additionally the ones found through the first stage of data coding, were identified and examined.

Results

Categories of Search Results

We found 5526 search queries (0.026%, 5526/21,000,000) that included the keyword suicide. The 5526 search queries were generated by 1586 unique search terms and 1625 unique users (0.25%, 1625/657,000). The three most frequently used search terms were "suicide" (9.79%, 541/5526), "suicide girls" (5.34%, 295/5526), and "how to commit suicide" (2.17%, 120/5526). Of the 5526 search queries, 3392 (61.38% of 5526) clicked through to a resulting webpage.

Of 3392 search queries with a click-through, 1344 unique webpages were clicked through by 930 unique users using 908 unique search terms. Of the 1344 webpages, 30 webpages were found to be inaccessible during the study period (October to November 2011) and returned error messages such as "page could not be found", "forbidden access", or "unable to display page" resulting in 1314 accessible clicked-through webpages. The three most frequently used search terms were also "suicide" (7.61%, 257/3392), "suicide girls" (4.95%, 168/3392), and "how to commit suicide" (2.74%, 93/3392).

Each clicked-through webpage was classified into 11 categories:

1. Entertainment (eg, writing, pictures, music about suicide): Webpages of original writing, pictures, music, or other content that may contain the keyword "suicide" for the purpose of amusement but not necessarily related to the conscious act of killing oneself.
2. Scientific Information: Webpages containing general educational materials, suicide statistics, and research studies.
3. Community Resources: Peer support or non-governmental organization webpages related to suicide prevention.
4. Specific Case/Incident: Webpages related to specific cases of highly publicized suicides.
5. Specific Means: Webpages that explicitly provide details on the methods of suicide, except drug overdose (see below).
6. Overdose and Suicide: There was a noticeable number of webpages that specifically addressed drug overdose or drug-related suicide information, and hence, a separate category from the Specific Means category was created.
7. Religion and Suicide: Webpages with information on suicide from a religious perspective.
8. Physician-Assisted Suicide: Webpages that contained information on euthanasia and physician-assisted suicide.
9. Discussion Forum: Any webpages with an interface primarily for discussion between users.
10. Families and Friends of Suicide/Attempt: Webpages with information about funeral arrangements and resources for family/friends of suicides.
11. Afterlife: Webpages with information and opinions related to "life after death" from suicide as well as life after surviving a suicide attempt. This category was singled out from the "religion and suicide" category because users used search terms that specifically included "afterlife", and these searches may or may not involve religiosity issues.

Among the 1314 webpages, the proportions of the webpage categories are as follows: entertainment (writings, pictures, music about suicide) (30.14%, 396/1314), scientific information (18.26%, 240/1314), community resources (14.54%, 191/1314), specific case/incident (11.04%, 145/1314), specific means (10.65%, 140/1314), overdose and suicide (4.72%, 62/1314), religion and suicide (4.11%, 54/1314), physician-assisted suicide (2.51%, 33/1314), discussion forum (2.28%, 30/1314), families and friends of suicide/attempt (1.60%, 21/1314), and afterlife (0.91%, 12/1314).

From the search queries, 78.01% of the traffic (2646/3392) went to a webpage link belonging to first page of the search results, 9.17% (311/3392) of the traffic went to the second page of search results, and 10.23% (347/3392) went to the third to fifth page of search results, while only 2.59% (88/3392) trickled to the sixth page or onwards in the search results. These results are generally consistent with current search engine data where links that are higher in the search engine results page, especially on the first page of the search results, tend to get majority of the traffic [29,30].

Clicked-Through Webpages and Search Terms Containing the Keyword “Suicide”

Table 1 shows the three most commonly used search terms for each webpage category, their usage frequencies, and the number of unique users for each search term. For webpages in the entertainment category, for example, the most commonly used search term that resulted in a click-through to a webpage was “suicide girls”, which was used 247 times (number of queries = 247) by 23 unique users.

Table 1. Top three search terms and their search frequencies in each category.

Category	Search term	Frequency	No. of unique users
1) Entertainment (eg, writings, pictures, and music about suicide)	suicide girls	247	23
	suicidegirls	90	11
	suicidegirls.com	46	8
	virgin suicides	46	2
2) Scientific information	suicide	69	49
	free essays on adolescent depression and suicide risks	28	1
	adolescent depression and suicide risks	25	1
3) Community resources	suicide	130	89
	suicide help	29	12
	suicide prevention	16	12
	teen suicide	16	11
4) Specific case/incident	famous suicide 1970 richard cunningham	28	1
	vachel lindsay suicide	16	1
	christopher lee anderson suicide	15	1
5) Specific means	how to commit suicide	70	21
	ways to commit suicide	36	8
	suicide	24	18
6) Overdose and suicide	how to commit suicide	17	13
	cymbalta and suicide	14	1
	suicide by overdose of kadian	8	1
	suicide by overdosing	8	1
7) Religion and suicide	how to commit suicide	13	10
	sermons preach for person who committed suicide	13	1
	suicide in the bible	11	3
8) Physician-assisted suicide	assisted suicide	46	19
	physician assisted suicide	13	8
	should america allow suicide euthanasia	8	1
9) Discussion forum	how to commit suicide	11	10
	Xxsuicide's xanga site	10	1
	suicideroadmap - myspace blog	9	8
10) Families and friends of suicide/attempt	suicide memorials	17	1
	survivors of suicide	8	3
	suicidememorialwall.com	4	1
11) Afterlife	suicide and the after life	9	1
	survivors of suicide	5	1
	after life suicide	4	1

When suicide was used as a single keyword search term (number of search queries = 256; 19.48%; 256/1314), users mostly clicked through to webpages that belong to the "community resources" (130 search queries), "scientific information" (69 search queries), and "specific means" categories (24 search queries). In contrast, when the search term "how to commit suicide" was used, the users most commonly clicked through to webpages in the "specific means" (70 queries), "overdose and suicide" (17 queries), and "religion and suicide" categories (13 queries).

The search terms most commonly used to access webpages in the "entertainment" category included the keywords in the names of a specific adult website and a novel cum movie. The "specific case/incident" category top search terms involved the keywords in the names of two famous individuals and a teenager who died from suicide; of note, each of these three search queries were contributed by one single user. Webpages in the "specific means" category were most commonly clicked through when users used the search terms "how to commit suicide", "ways to commit suicide", and "suicide".

Table 2 shows the top three most clicked webpages of each category. The webpages with more than 40 click-throughs during the data collection period included "http://www.suicidegirls.com" (entertainment), "http://suicidegirls.com" (entertainment), "http://metanoia.org" (community resource), "http://www.satanservice.org" (specific means), and "http://www.cdc.gov" (scientific information). In the entertainment category, the two most clicked webpages led to the same web domain, which was an adult website. Wikipedia, an online information resource (Figure 1), followed by an automobile parts branded e-store ("http://www.suicidedoors.com") were also in the most clicked webpages for the entertainment category. For the specific case/incident category, Wikipedia came up on top again with the most click-throughs, in addition to a news webpage and a social networking webpage.

Figure 1. Suicide page on Wikipedia.

Suicide

From Wikipedia, the free encyclopedia

For other uses, see Suicide (disambiguation).

Suicide (Latin *suicidium*, from *sui caedere*, "to kill oneself") is the act of intentionally causing one's own death. Suicide is often committed out of despair, the cause of which can be attributed to a mental disorder such as depression, bipolar disorder, schizophrenia, alcoholism, or drug abuse.[1] Stress factors such as financial difficulties or troubles with interpersonal relationships often play a significant role.[2]

Over one million people die by suicide every year. The World Health Organization (WHO) estimates that it is the 13th leading cause of death worldwide[3] and the National

Suicide

Classification and external resources

The Suicide by Édouard Manet 1877–1881

ICD-10	X60–X84
ICD-9	E950
MedlinePlus	001554
eMedicine	article/288598
MeSH	*F01.145.126.980.875*

Table 2. Top three most clicked-through webpages of each category, their click-through frequencies, and the number of unique users.

Category	Webpage	Frequency	No. of unique users
1) Entertainment (writings, pictures, music, about suicide)	http://www.suicidegirls.com	153	76
	http://suicidegirls.com	78	59
	http://en.wikipedia.org	21	18
	http://www.suicidedoors.com	21	16
2) Scientific information	http://www.cdc.gov	43	35
	http://www.psycom.net	25	23
	http://www.ncbi.nlm.nih.gov	15	13
3) Community resources	http://www.metanoia.org	86	81
	http://kidshealth.org	20	17
	http://www.afsp.org	16	14
	http://www.focusas.com	16	10
4) Specific case/incident	http://en.wikipedia.org	18	13
	http://news.bbc.co.uk	7	5
	http://profile.myspace.com	7	6
5) Specific means	http://www.satanservice.org	49	39
	http://www.mouchette.org	21	18
	http://extremeriver.org	14	13
6) Overdose and suicide	http://www.a1b2c3.com	31	23
	http://www.ncbi.nlm.nih.gov	7	5
	http://www.everything2.com	5	5
7) Religion and suicide	http://www.religioustolerance.org	16	13
	http://www.believers.org	13	11
	http://en.wikipedia.org	7	5
8) Physician-assisted suicide	http://www.assistedsuicide.org	13	11
	http://www.religioustolerance.org	10	9
	http://www.amsa.org	5	5
	http://www.euthanasia.com	5	5
9) Discussion forum	http://www.xanga.com	14	2
	http://www.zenhex.com	13	12
	http://lifeflame.6.forumer.com	5	5
10) Families and friends of suicide/attempt	http://www.suicidememorialwall.com	9	4
	http://www.survivorsofsuicide.com	7	6
	http://www.parentsofsuicide.com	5	4
11) Afterlife	http://www.near-death.com	5	3
	http://samvak.tripod.com	3	1
	http://www.survivorsofsuicide.com	3	2

Search Terms Used to Find Pro-suicide Websites

Among the 1314 accessed webpages, we could identify only two clicked-through websites that were considered pro-suicide. They are "http://www.suicidemethods.net" and "http://www.churchofeuthanasia.org". The two websites not only provide information on methods of suicide but also portray a positive attitude towards the choice of using suicide as a way to ease pain. Table 3 shows the searches terms that were used to access the two websites (ie, "how to kill yourself", "best way to commit suicide", and "euthanasia"), the frequency of their usage, and the number of unique users for each search term.

We found that the commonly used search terms included “pictures of murder by strangulation” and “photo of a severe burn” for “http://www.suicidemethods.net” and “committing suicide with a gas oven”, and “hairless goat” for “http://www.churchofeuthanasia.org”. It is noteworthy that 23 and 14 unique users visited “http://www.suicidemethods.net” and “http://www.churchofeuthanasia.org” respectively during the data collection period. It is important to view these numbers in context, especially given the fact that a single unique user contributed to the entire search query frequency of both “committing suicide with a gas oven” and “methods to commit suicide”, while the top two of search query frequency for “http://www.suicidemethods.net” were a result of only two unique users.

Table 3. Search terms used to access the two pro-suicide websites.

Website	Search term	Frequency	No. of unique users
http://www.churchofeuthanasia.org	committing suicide with a gas oven	27	1
	hairless goat	22	1
	how to kill yourself	18	3
	methods to commit suicide	11	1
	euthanasia	10	2
	best way to commit suicide	5	1
	butcher pig	4	1
	committing suicide	4	1
	how to die painlessly	3	1
	the church of euthanasia	1	1
	butchering the human carcass for human consumption	1	1
	dog sex kkh	1	1
http://www.suicidemethods.net	pictures of murder by strangulation	26	1
	photo of a severe burns	22	1
	pro choice suicide	7	1
	suicidal websites	6	1
	suicide murder pics	6	1
	suicide pics	6	1
	suicide methods	5	2
	asphyxia pics	4	1
	gory pictures	3	3
	gory photos	2	2
	homicide pictures	2	2
	pictures of suicides	2	2
	suicide	2	1
	bloody suicide jump pictures	1	1
	cut wrist pictures	1	1
	gory autopsy pictures	1	1
	hanging suicide	1	1
	shotgun wounds	1	1
	suicide attempt	1	1
	suicide and hell	1	1
	suicide photos	1	1

Discussion

This is a retrospective observational study aimed to investigate the naturalistic web searching behavior on a search engine by online users looking for content using suicide-related keywords. Specifically, we examined users' search terms with the keyword "suicide" and investigated the types of webpages they had accessed in the search results of AOL's search engine, one of the major search engines in the United States during the data collection period. As such, the database used in this study might be dated, confined to an English-dominated search engine, and mainly limited to North America; nonetheless, this is the most comprehensive database of its kind that is freely available in the public domain for exploration. Due to online privacy developments, no other comparable publicly available datasets exist. Although using Google's search trend tool, Google Trends, may provide some data on searching behaviors, comparisons would be difficult on many levels, including sampling ambiguity, lack of information on absolute search numbers, normalization of search results, lack of details on scaling of search results, and inability to access fine-grained searches.

Despite being a older dataset, especially in the context of the Internet where content and usage behavior is constantly changing, the findings still have relevance because, unlike prior studies that took a more computational perspective to suicide-related Internet searches, we used a naturalistic approach by looking at the users' actual search behaviors. Rather than merely looking at what is available for our perusal on the Internet, which is subject to constant change, we examined the subsequent behavior with returned results, which seems relatively consistent even over time [29,30]. Furthermore, the detailed results of the search terms that we have been able to access through this dataset are not currently available anywhere since search trend tools are not able to drill down to this level of detail, and therefore, this may be the only glimpse into detailed search behavior of online users.

Utilizing this type of dataset, however, means that although users' search result click-throughs were logged, we do not have a record of the search results webpages they could see and potentially access. Also, users' behavior after accessing the webpages cannot be examined. Nevertheless, it is possible to identify search trends, patterns, and popular click paths that allow us to gain insight into users' searching goals and intent as well as pages that had high visibility based on the keywords used in the search queries.

Presumably there are many existing websites that contain information about suicide, making the Internet's role in suicides a point of concern. Our findings highlight three important findings on Internet users' naturalistic web search behaviors that may challenge the underlying assumptions behind past studies suggesting that the Internet is intrinsically harmful because pro-suicide information can be found easily using suicide-related search terms. First, despite previous studies finding that pro-suicide and how-to-commit-suicide websites are easily accessible through search engines [15,21-23], our study showed that for AOL searches with suicide-related search terms, users generally accessed webpages in the search results that provided entertainment, scientific information, news, and resource information. About 10% of webpages accessed included information about specific means of suicide, or about 15% if overdose and suicide webpages are included. These preliminary findings could be interpreted as an indication of users' search intent when they use suicide-related search terms. Second, only a very small proportion of users visited websites in the search results that were traditionally seen as potentially negative or harmful websites and among the clicked-through webpages, we could identify only two websites that may be considered as pro-suicide, showing detailed information of how to complete suicide and encouraging viewers to use suicide as a problem-solving strategy. Third, further investigation revealed that the majority of users did not use search terms that contained the keyword "suicide" to gain access to the two pro-suicide websites. More interestingly, according to the search terms that were used to access the pro-suicide websites, users may have been searching for gory images of unnatural deaths rather than for descriptive information on ways to commit suicide.

The Internet and Suicide Prevention

In traditional media, newspapers, radio, and television were the major sources for information and their consumers were perceived as passive receivers of information. Information portrayed in traditional media was mainly contributed by professional journalists, and hence, regulation of that information was relatively straightforward. In terms of suicide information in traditional media, since information about suicide cases are presented and received passively by the masses, media reporting guidelines on suicide information developed for professionals has been considered as a core prevention strategy [31]. With the development of the Internet, we have stepped into the information age, which has revolutionized the ways we consume and produce information and transformed how we organize knowledge [32].

In the Web 2.0 era, the new media is interactive. People become proactive consumers and producers of information. Not only can people search for information and contribute, messages and information can also be widely and easily shared among friends via social media and networks. The ease and low cost of creating and copying content on the Internet contributed to its exponential growth and now, the amount of information that can be found on the Internet can be considered as almost infinite. In 2004, Google's whole data storage was approximately five petabytes, which is the quantity of data a thousand times larger than the Library of Congress's print collection [33]. According to a recent survey, there are 346,004,403 websites on the Internet as of June 2011 [34], and Google announced on July 2008 that it had processed over one trillion unique URLs [35], while Bing estimated that there were one trillion pages of content in 2009 [36]. Internet users, who can now also be contributors of information, do not have any obligations to follow reporting guidelines. Furthermore, content on the Internet can be easily copied, re-posted, cached, and reproduced in mirror-websites, and blocked websites can be bypassed easily through encrypted connections. A good example is the WikiLeaks site "http://en.wikipedia.org/wiki/WikiLeaks", which had many mirror websites created when authorities tried to shut down and remove all the content from original website. The properties of this new medium make restricting access to particular websites difficult and impractical. Although the exact number of pro-suicide websites on the Internet is unknown, it is close to impossible to remove such websites from the Internet, regardless of whether they are easy or difficult to access.

Actual Behavior and Intent of Suicide-Related Searches Online

Internet users use the Internet for various reasons. Segev and Ahituv conducted a cross-national analysis of popular search queries used in Google and Yahoo! over a 24-month period from January 2004 to December 2005 and found that there are cultural differences related to Internet usage [10]. They found that in many English-speaking and Western countries, the most popular searches were related to entertainment. According to the Pew Internet and American Life Project, information searches and email are the two major online activities among adult Internet users [37], and communication and entertainment are the most common ones among young people [38]. Accordingly, the "uses and gratifications theory" [39], a well-known theory in the field of media studies, states that the information people choose to access is largely dependent on their media consumption needs and desire. In other words, although negative information is easily accessible through new media, whether people access it largely depends on what they are looking for.

Our findings have provided a glimpse into the information consumption preferences of online users and corroborates prior studies on people's Internet usage. Even for a domain-specific topic, suicide, our study generated consistent findings that most search engine users in the United States used suicide-related queries to search for websites that belong to the entertainment category.

Although pro-suicide websites and websites that contain detailed information on suicide methods can be accessed easily, the majority of search engine users, at least in the United States, did not access them using suicide-related queries. A number of keywords used to access pro-suicide webpages were instead related to violent or bloody pictures including "gory pictures", "gory photos", and "homicide pictures". With these queries, we suspect that those who accessed the websites might not have an intent to die. Westerlund [40] has suggested that the fascination with pro-suicide content, particularly the morbid and violent descriptions may be a manifestation of a meaningful process in which the producer or consumer of this content is building an identity, acting out aggressive impulses, or even rebelling against the dominant culture. An interesting future study would be to explore the motivations behind why non-suicidal individuals access suicide-related information. The AOL data are very much based on a Western population and according to Segev and Ahituv's findings [10], we do not know whether people in other countries behave similarly or differently. Recently developed free tools, like Google Trends and Google Insights for Search, may help to conduct a cross-national comparison study on this topic. With that information, however, we are still far from understanding how Internet users process information collected from the Internet and whether pro-suicide information leads to suicidal behaviors.

In-depth interviews with individuals with nearly lethal suicidal behavior may help to understand this phenomenon. In Biddle et al's study [25], 22 individuals who survived nearly fatal suicide attempts were interviewed on the information sources that informed their choice of suicide method. They found that about 36% of the attempters' choices of suicide method were found through the Internet. Interestingly, according to the transcripts cited in their paper, some attempters were "inspired" by online footage of Saddam Hussein's execution or Wikipedia, which contains detailed information on suicide methods despite being a general information website. They wrote "the sites accessed by these respondents tended to be those containing professional information and resources (including online chemists), general knowledge sites (eg, Wikipedia), or news sites, including BBC news. Specific 'suicide sites' were accessed less frequently…" (p.705 of [25]). This information is essential to inform further development of suicide prevention work on the Internet.

In our study, we found that a small group of individuals intentionally searched for information about ways to complete suicide and an even smaller group of individuals were eager to find discussion forums using suicide-related keywords. A few individuals used queries like "suicide by overdose of XX" or "suicide by XX" to get access to webpages that provide information on prescribed drugs and suicide. In contrast, websites that include information on suicide means were accessed by larger numbers of unique users. One individual used a specific discussion forum address as a query ("xxsuicide's xanga site" as shown in the table) to get access to that particular webpage 10 times.

There is a lack of research on suicide pacts arranged online, and so it is difficult to predict the major consequence of a large number of discussants getting access to this sort of discussion forum. On one hand, there have been some documented instances of completed suicide pacts in Japan [18], and on the other hand, there is evidence for the positive effects of joining an online community, like Facebook, which may enhance ones' positive affective state and also increase ones' social support [41].

What Can Be Done to Minimize the Negative Impact of the Internet on Suicide?

Current efforts to reduce the potential negative impact of the Internet on suicide include remedial actions like removing and blocking entries of harmful websites by using Internet filters, constant monitoring of potential harmful websites by "web-based police officers" as implemented in Japan and Korea, and displaying helpline resources information while users key in suicide-related queries. These strategies are in place in an attempt to either shield suicidal and vulnerable individuals from suggestive material or abort their suicidal wishes by diverting their attention to help resources. There are several challenges that these strategies face. First, these strategies do not observe the way that information is created, found, and accessed online. Accordingly then, restricting or removing potentially harmful websites is an arduous if not impossible task. Furthermore, cyber regulations are extremely difficult to enforce since the

Internet is not clearly under any jurisdiction, and identification of a physical user can be elusive.

Instead, it is important to further investigate the natural online behavior especially of vulnerable individuals and integrate that information with the way that search engines provide information. With the Internet inundated with information, the major search engines, ie, Google, Bing, and Yahoo!, have complex algorithms to return relevant search results to users. Consistent with the uses and gratification theory, relevance in a search engine is defined by variables such as user popularity, credible web domains, and high-quality content. As such, well-made and informative websites will tend to rank higher in search results and also be visited more often. Consequently, instead of the labor-intensive strategy of finding, sorting, and removing potentially harmful websites, more resources and effort should be spent on developing high-quality, informative, interactive, and user-friendly websites that maximize the likelihood of ranking highly in search results and therefore, being found and accessed.

Also, more recent studies find that the Internet provides more positive than negative effects on the vulnerable, especially in enhancing the social support of isolated individuals though social networking sites [11,42]. Mental health providers and researchers should make use of Information and Communication Technology (ICT) in strengthening their traditional practices in a fashion that could not be achieved in pre-digital times.

Conclusion

Our research presented the naturalistic search behavior of search engine subscribers using any search terms with the keyword "suicide". Currently we know that cybersuicide or Internet suicide pacts represent a small fraction of overall suicides. While those attempting suicide have searched for suicide-related information on the Internet before completing the act, we also know that the Internet can provide a unique platform to reach those individuals previously inaccessible, and web-based psychotherapies are producing promising results in helping depressed individuals, with the acknowledgment of the potential digital divide phenomenon. As such, it can be concluded that we should neither underestimate nor overestimate the potential negative impact of the Internet on suicide.

However, the threshold between offline and online activities is quickly receding and in the near future will soon disappear. It is uncertain how much of today's searching, and even learning, behavior will remain unchanged. There are plenty of daily examples that show how much of our learning behavior has transformed. *"Siri" (*Speech Interpretation and Recognition Interface) is an intelligent personal assistant and knowledge navigator that functions as an application for Apple's operating system and answers questions, makes recommendations, and performs actions by delegating requests to a set of web services to iPhone users. Google Images, a searching tool for images on the Internet, may help Google users to search for visual information using image or pictures. These are just a few examples. Suicide research and prevention strategies must keep up with the emerging trend and pace of ICT in order to prevent more unnecessary tragic deaths.

Acknowledgments

We thank the anonymous reviewers for their helpful suggestions. The publication of this article was partly supported by the General Research Fund (HKU 756211H granted to PW).

Conflicts of Interest

None declared.

References

1. International Telecommunication Union. The World in ICT Facts and Figures. 2011. URL: http://www.itu.int/ITU-D/ict/facts/2011/material/ICTFactsFigures2011.pdf [accessed 2012-05-21] [WebCite Cache ID 67qE7fXIW]
2. Hagihara A, Miyazaki S, Abe T. Internet suicide searches and the incidence of suicide in young people in Japan. Eur Arch Psychiatry Clin Neurosci 2012 Feb;262(1):39-46. [doi: 10.1007/s00406-011-0212-8] [Medline: 21505949]
3. Boyce N. Pilots of the future: suicide prevention and the internet. Lancet 2010 Dec 4;376(9756):1889-1890. [Medline: 21155079]
4. 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]
5. Eysenbach G. Infodemiology and infoveillance tracking online health information and cyberbehavior for public health. Am J Prev Med 2011 May;40(5 Suppl 2):S154-S158. [doi: 10.1016/j.amepre.2011.02.006] [Medline: 21521589]
6. Eysenbach G. Infodemiology: tracking flu-related searches on the web for syndromic surveillance. AMIA Annu Symp Proc 2006:244-248 [FREE Full text] [Medline: 17238340]
7. Ayers JW, Althouse BM, Allem JP, Ford DE, Ribisl KM, Cohen JE. A novel evaluation of World No Tobacco day in Latin America. J Med Internet Res 2012;14(3):e77 [FREE Full text] [doi: 10.2196/jmir.2148] [Medline: 22634568]

8. Hill S, Mao J, Ungar L, Hennessy S, Leonard CE, Holmes J. Natural supplements for H1N1 influenza: retrospective observational infodemiology study of information and search activity on the Internet. J Med Internet Res 2011;13(2):e36 [FREE Full text] [doi: 10.2196/jmir.1722] [Medline: 21558062]
9. Dobson R. Internet sites may encourage suicide. BMJ 1999 Aug 7;319(7206):337 [FREE Full text] [Medline: 10435946]
10. Segev E, Ahituv N. Popular searches in Google and Yahoo!: A "digital divide" in information uses? The Information Society 2010 Jan;26(1):17-37. [doi: 10.1080/01972240903423477]
11. Ruder TD, Hatch GM, Ampanozi G, Thali MJ, Fischer N. Suicide announcement on Facebook. Crisis 2011;32(5):280-282. [doi: 10.1027/0227-5910/a000086] [Medline: 21940257]
12. Page A, Chang SS, Gunnell D. Surveillance of Australian suicidal behaviour using the internet? Aust N Z J Psychiatry 2011 Dec;45(12):1020-1022. [doi: 10.3109/00048674.2011.623660] [Medline: 22034830]
13. Messias E, Castro J, Saini A, Usman M, Peeples D. Sadness, suicide, and their association with video game and internet overuse among teens: results from the youth risk behavior survey 2007 and 2009. Suicide Life Threat Behav 2011 Jun;41(3):307-315. [doi: 10.1111/j.1943-278X.2011.00030.x] [Medline: 21463355]
14. Durkee T, Hadlaczky G, Westerlund M, Carli V. Internet pathways in suicidality: a review of the evidence. Int J Environ Res Public Health 2011 Oct;8(10):3938-3952 [FREE Full text] [doi: 10.3390/ijerph8103938] [Medline: 22073021]
15. Thompson S. The Internet and its potential influence on suicide. Psychiatric Bulletin 1999 Aug;23(8):449-451. [doi: 10.1192/pb.23.8.449]
16. Ozawa-de Silva C. Too lonely to die alone: internet suicide pacts and existential suffering in Japan. Cult Med Psychiatry 2008 Dec;32(4):516-551. [doi: 10.1007/s11013-008-9108-0] [Medline: 18800195]
17. Lee DT, Chan KP, Yip PS. Charcoal burning is also popular for suicide pacts made on the internet. BMJ 2005 Mar 12;330(7491):602 [FREE Full text] [doi: 10.1136/bmj.330.7491.602-b] [Medline: 15761009]
18. Rajagopal S. Suicide pacts and the internet. BMJ 2004 Dec 4;329(7478):1298-1299 [FREE Full text] [doi: 10.1136/bmj.329.7478.1298] [Medline: 15576715]
19. Becker K, Schmidt MH. Internet chat rooms and suicide. J Am Acad Child Adolesc Psychiatry 2004 Mar;43(3):246-247. [Medline: 15076254]
20. Kemp CG, Collings SC. Hyperlinked suicide: assessing the prominence and accessibility of suicide websites. Crisis 2011;32(3):143-151. [doi: 10.1027/0227-5910/a000068] [Medline: 21616763]
21. Biddle L, Donovan J, Hawton K, Kapur N, Gunnell D. Suicide and the internet. BMJ 2008 Apr 12;336(7648):800-802 [FREE Full text] [doi: 10.1136/bmj.39525.442674.AD] [Medline: 18403541]
22. Cheng Q, Fu KW, Yip PS. A comparative study of online suicide-related information in Chinese and English. J Clin Psychiatry 2011 Mar;72(3):313-319. [doi: 10.4088/JCP.09m05440blu] [Medline: 20868633]
23. Recupero PR, Harms SE, Noble JM. Googling suicide: surfing for suicide information on the Internet. J Clin Psychiatry 2008 Jun;69(6):878-888. [Medline: 18494533]
24. Grohol JM. Suicide and the internet: Study misses internet's greater collection of support websites. BMJ 2008 Apr 26;336(7650):905-906 [FREE Full text] [doi: 10.1136/bmj.39556.435000.80] [Medline: 18436921]
25. Biddle L, Gunnell D, Owen-Smith A, Potokar J, Longson D, Hawton K, et al. Information sources used by the suicidal to inform choice of method. J Affect Disord 2012 Feb;136(3):702-709. [doi: 10.1016/j.jad.2011.10.004] [Medline: 22093678]
26. Fu KW. What do people want to know about depression? Analyzing a representative web search queries sample. 2009 Mar Presented at: Emerging Mode of Communication: Technology Enhanced Interaction; March 2009; Hong Kong SAR, China p. 26-27.
27. comScore Networks. Google sees its U.S. search market share jump a full point in May, marking tenth consecutive monthly gain. 2006. URL: http://ir.comscore.com/releasedetail.cfm?ReleaseID=245438 [accessed 2012-05-22] [WebCite Cache ID 67qJOatuW]
28. Pass G, Chowdhury A, Torgeson C. A picture of search. In: Proceedings of The First International Conference on Scalable Information Systems. New York, NY: ACM; 2006 Presented at: First International Conference on Scalable Information Systems; May 29-June 1, 2006; Hong Kong SAR, China. [doi: 10.1145/1146847.1146848]
29. McGee M. Top Google Ranking Gets Twice The Traffic Of #2 Ranking: Chitika. 2010. URL: http://searchengineland.com/value-of-top-google-ranking-42904 [accessed 2012-05-22] [WebCite Cache ID 67qPyoGUi]
30. Goodwin D. Top Google Result Gets 36. 2011. URL: http://searchenginewatch.com/article/2049695/Top-Google-Result-Gets-36.4-of-Clicks-Study [accessed 2012-05-22] [WebCite Cache ID 67qQDcXvW]
31. World Health Organization. Preventing suicide: A resource for media professionals. 2008. URL: http://www.who.int/mental_health/prevention/suicide/resource_media.pdf [accessed 2012-05-22] [WebCite Cache ID 67qbYSHux]
32. Floridi L. Internet: Which Future for Organized Knowledge, Frankenstein or Pygmalion? International Journal of Human-Computer Studies 1995 Aug;43(2):261-274. [doi: 10.1006/ijhc.1995.1044]
33. Mellor C. Google's storage strategy - not a SAN in sight. 2004 URL: http://features.techworld.com/storage/467/googles-storage-strategy/ [accessed 2012-05-22] [WebCite Cache ID 67qRoPZv9]
34. How many websites are there on the internet?. 2011. URL: http://howmanyarethere.net/how-many-websites-are-there-on-the-internet/ [accessed 2012-05-22] [WebCite Cache ID 67qSXfBdi]

35. Alpert J, Hajaj N. We knew the web was big. 2008 URL: http://googleblog.blogspot.com/2008/07/we-knew-web-was-big.html [accessed 2012-05-22] [WebCite Cache ID 67qSl14W9]
36. Nadella S. User Needs, Features and the Science behind Bing. 2009. URL: http://www.bing.com/community/site_blogs/b/search/archive/2009/06/01/user-needs-features-and-the-science-behind-bing.aspx [accessed 2012-05-22] [WebCite Cache ID 67qSvkijI]
37. Zickuhr K, Smith A. Digital differences. 2012. URL: http://pewinternet.org/~/media//Files/Reports/2012/PIP_Digital_differences_041312.pdf [accessed 2012-05-22] [WebCite Cache ID 67qbFJTNp]
38. Moreno MA, Jelenchicka L, Koffb R, Eikoffc J, Diermyerd C, Christakise DA. Internet use and multitasking among older adolescents: An experience sampling approach. Computers in Human Behavior 2012 Feb;28(4):1097-1102. [doi: 10.1016/j.chb.2012.01.016]
39. Katz E, Blumler JG, Gurevitch M. Utilization of mass communication by the individual. In: Blumler JG, editor. The uses of mass communications: Current perspectives on gratifications research. Beverly Hills, CA: Sage Publications; 1974:19-32.
40. Westerlund M. The production of pro-suicide content on the Internet: a counter-discourse activity. New Media & Society 2011.
41. Mauri M, Cipresso P, Balgera A, Villamira M, Riva G. Why is Facebook so successful? Psychophysiological measures describe a core flow state while using Facebook. Cyberpsychol Behav Soc Netw 2011 Dec;14(12):723-731. [doi: 10.1089/cyber.2010.0377] [Medline: 21879884]
42. Silenzio VM, Duberstein PR, Tang W, Lu N, Tu X, Homan CM. Connecting the invisible dots: reaching lesbian, gay, and bisexual adolescents and young adults at risk for suicide through online social networks. Soc Sci Med 2009 Aug;69(3):469-474 [FREE Full text] [doi: 10.1016/j.socscimed.2009.05.029] [Medline: 19540641]

Edited by G Eysenbach; submitted 23.05.12; peer-reviewed by M Westerlund; comments to author 25.08.12; revised version received 30.08.12; accepted 26.09.12; published 11.01.13.

Please cite as:
Wong PWC, Fu KW, Yau RSP, Ma HHM, Law YW, Chang SS, Yip PSF
Accessing Suicide-Related Information on the Internet: A Retrospective Observational Study of Search Behavior
J Med Internet Res 2013;15(1):e3
URL: http://www.jmir.org/2013/1/e3/
doi:10.2196/jmir.2181
PMID:23305632

Original Paper

Health Professionals' Expectations Versus Experiences of Internet-Based Telemonitoring: Survey Among Heart Failure Clinics

Arjen E de Vries[1], RN, MA; Martje H.L van der Wal[1], RN, PhD; Maurice M.W Nieuwenhuis[1], PhD; Richard M de Jong[1], MD, PhD; Rene B van Dijk[2], MD, PhD; Tiny Jaarsma[3], RN, PhD; Hans L Hillege[1], MD, PhD

[1]University Medical Center Groningen, Cardiology, Groningen, Netherlands

[2]Martini Hospital, Cardiology, Groningen, Netherlands

[3]Linkoping University, Department of Social and Welfare Studies, Norrkoping, Sweden

Corresponding Author:
Arjen E de Vries, RN, MA
University Medical Center Groningen
Cardiology
Hanzeplein 1
Groningen, 9700 RB
Netherlands
Phone: 31 503612094
Fax: 31 503618062
Email: a.e.de.vries@thorax.umcg.nl

Abstract

Background: Although telemonitoring is increasingly used in heart failure care, data on expectations, experiences, and organizational implications concerning telemonitoring are rarely addressed, and the optimal profile of patients who can benefit from telemonitoring has yet to be defined.

Objective: To assess the actual status of use of telemonitoring and to describe the expectations, experiences, and organizational aspects involved in working with telemonitoring in heart failure in the Netherlands.

Methods: In collaboration with the Netherlands Organization for Applied Scientific Research (TNO), a 19-item survey was sent to all outpatient heart failure clinics in the Netherlands, addressed to cardiologists and heart failure nurses working in the clinics.

Results: Of the 109 heart failure clinics who received a survey, 86 clinics responded (79%). In total, 31 out of 86 (36%) heart failure clinics were using telemonitoring and 12 heart failure clinics (14%) planned to use telemonitoring within one year. The number of heart failure patients receiving telemonitoring generally varied between 10 and 50; although in two clinics more than 75 patients used telemonitoring. The main goals for using telemonitoring are “monitoring physical condition”, “monitoring signs of deterioration” (n=39, 91%), “monitoring treatment” (n=32, 74%), “adjusting medication” (n=24, 56%), and “educating patients” (n=33, 77%). Most patients using telemonitoring were in the New York Heart Association (NYHA) functional classes II (n=19, 61%) and III (n=27, 87%) and were offered the use of the telemonitoring system “as long as needed” or without a time limit. However, the expectations of the use of telemonitoring were not met after implementation. Eight of the 11 items about expectations versus experiences were significantly decreased ($P<.001$). Health care professionals experienced the most changes related to the use of telemonitoring in their work, in particular with respect to “keeping up with current development” (before 7.2, after 6.8, $P=.15$), “being innovative” (before 7.0, after 6.1, $P=.003$), and “better guideline adherence” (before 6.3, after 5.3, $P=.005$). Strikingly, 20 out of 31 heart failure clinics stated that they were considering using a different telemonitoring system than the system used at the time.

Conclusions: One third of all heart failure clinics surveyed were using telemonitoring as part of their care without any transparent, predefined criteria of user requirements. Prior expectations of telemonitoring were not reflected in actual experiences, possibly leading to disappointment.

(J Med Internet Res 2013;15(1):e4) doi:10.2196/jmir.2161

KEYWORDS

Telemonitoring, Telemedicine, Remote monitoring, Internet, Heart failure, Heart failure management, Disease management,

Introduction

Telemonitoring in heart failure care is used to monitor patients' symptoms at home and to guide patients in taking action in case of deterioration. Telemonitoring is considered a promising new intervention for heart failure patients, and a study on the use, perceptions, and experiences has been published recently [1,2]. However, current evidence regarding the effectiveness of telemonitoring in the care of heart failure patients is conflicting [3]. There are many definitions used for telemonitoring, but the core principle does not generally differ. A commonly used international definition is "the remote monitoring of patients, including the use of audio, video, and other telecommunications and electronic information processing technologies to monitor patient status at a distance" [4]. In the Netherlands, the most used definition is that telemonitoring includes the measurement, monitoring, collecting, and transfer of clinical data concerning the health status of a patient in his or her home environment, using information and communication technology. Initial studies showed that remote monitoring of heart failure patients reduced hospitalization and mortality rates [5-8]. However, recent studies performed on a larger scale did not confirm these findings [9,10]. Questions remain regarding the optimal patient profile for using telemonitoring, the technical aspects of the telemonitoring systems, the intensity and frequency of providing data, and the cost-effectiveness of the various telemonitoring systems used [11,12]. Furthermore, expectations and consequences of telemonitoring for the organization of care, logistic processes, and the work of health care providers are rarely studied, and thus unclear. However, these aspects of telemonitoring are vital for the consideration and acceptance of these systems in future practice [13].

Despite the inconclusive evidence for the use of telemonitoring in heart failure, telemonitoring is considered to be a promising development, [7] and there are increasing efforts to introduce telemonitoring in outpatient heart failure clinics. In some countries, including the Netherlands, health care insurance companies reimburse telemonitoring for heart failure patients. The present study was designed to assess the perspectives and expectations for both heart failure nurses and cardiologists working in a heart failure team with telemonitoring.

To this end, the following research questions were posed: 1) What are the perceptions and expectations of cardiologists and heart failure nurses with respect to the implementation of telemonitoring in heart failure patients? and 2) What are their experiences with the implementation of telemonitoring? In this study, we did not focus on possible differences between heart failure nurses and cardiologist in their perceptions of working with telemonitoring.

Methods

Participants

Participants in the study consisted of cardiologists and heart failure nurses working in heart failure outpatient clinics in the Netherlands. Out of all 118 Dutch heart failure clinics, 109 clinics received a questionnaire in March 2011, addressed to the cardiologists and heart failure nurses working in the heart failure outpatient clinic. Nine heart failure clinics were excluded and did not receive a questionnaire due to their participation in the IN TOUCH study, a study evaluating the added value of information and communication technology-guided disease management combined with telemonitoring for heart failure patients [14]. Participants were requested to return the questionnaire within 12 weeks. We sent out two reminders.

Instrument

In collaboration with the Netherlands Organization for Applied Scientific Research (TNO), a 19-item questionnaire on telemonitoring was specifically developed for this study, based on the two research questions. For this questionnaire we defined telemonitoring as: "The remote, Internet-based monitoring and mentoring of heart failure patients on weight, blood pressure, heart rate, and signs and symptoms that disclose the actual condition of the heart failure patient. The devices are used by the patients in their own home environment and the generated data are transferred by the Internet". The use of telemonitoring by means of telephone, telephone support, telephone follow-up, or by means of implantable devices was not included in this study because our focus was to investigate expectations and experiences of using telemonitoring devices that required an active user interaction (eg, direct handling of deviated values, generated alerts, and complaints). The technology and handling for users between implanted devices and external devices, such as weight scales and/or blood pressure measurements, are essentially different. Based on the research questions, items for the questionnaire were developed with the input of 10 cardiologists and 10 heart failure nurses, resulting in a questionnaire consisting of 3 domains: 1) availability of telemonitoring, 2) experiences with telemonitoring, and 3) organization of telemonitoring. The questionnaire consisted of both multiple choice and "agree/disagree" questions. For data regarding the motivation for and importance of using telemonitoring, as well as the experiences with using telemonitoring, we asked respondents to rate 11 items on a 10-point scale. On this scale, 0 counted as "not important"' and 10 as "very important".

These 11 items were based on practical considerations related to the start-up of telemonitoring. Aside from addressing the practical considerations of health care workers in our study, these same 11 items are frequently used by sales representatives to convince future users of the added value of working with telemonitoring. The 11 different items could be combined into 3 groups: 1) direct patient care (better self-management, improving quality of care, and reduction of (re) admission); 2) telemonitoring system–related aspects (current development, innovation, and better guideline adherence); and 3) organizational aspects (treating more patients, fulfilling hospital policy, reducing workload, lowering heart failure related costs, and fulfilling health care insurance policy).

Validation Process of the Questionnaire

To test the questionnaire, a group of 30 pilot responders, representing the future research population, completed the questionnaire. Internal consistency (Cronbach alpha) of the questionnaire in the current sample was .85. This parameter measures the reliability of the scale. A set of questionnaire items with a reliability of .70 or higher is considered acceptable. Face validity (10 cardiologists, 10 heart failure nurses) was assessed by analyzing the feedback received on the total questionnaire.

Statistical Analysis

Descriptive statistics were used to present the data. For some parts of the analysis, we subdivided the respondents into current telemonitoring users (n=31) and intended telemonitoring users (n=12), because some research questions are related to actual experiences of working with telemonitoring and other are more exploratory (eg, which patients do you think are suitable for applying telemonitoring?). Paired samples *t* tests were used to examine possible differences between expectations of and experiences with using telemonitoring. Analyses were performed using PASW, version 18.0 for Windows.

Results

Basic Characteristics of the Study Population

Of the 109 heart failure clinics who received a survey, 86 clinics responded (79%). Their responses were included in the analysis. Respondents had a mean age of 48 ± 8 years, and 68% were female. The mean years of work experience in the current position was 14 ± 9 years, and the respondents worked with heart failure patients for an average of 19 ± 10 hours a week. Of the 86 responding clinics, 31 reported using telemonitoring in their current patient care (36 %), and 12 clinics (14%) planned to use telemonitoring within one year. Further analysis was therefore restricted to the clinics that actually used telemonitoring and those that planned to use telemonitoring within one year (total n=43).

Availability of Telemonitoring

The three systems most frequently used for telemonitoring were commercially available systems (Motiva, Health Buddy, and IPT Telemedicine [15-17]), and one clinic had developed its own telemonitoring system. The systems used in this study are generally similar to each other based on functionality. They transfer measurements generated at home and answers to questions to a health care environment via the Internet. The Health Buddy system differs, however, because it transfers the data directly to the health care provider instead of a data center. This means that the heart failure nurses are directly responsible for the handling of data and measurements. However, the consequence of directly receiving data and measurements is the need for a 24/7 shift of health care providers.

The feedback from the health care provider to the patient in all three systems is given by telephone. For the specific characteristics of the commercially available systems used in this study [18], see Table 1.

Table 1. Characteristics of the commercial available telemonitoring systems used in this study.

		Motiva	Health Buddy	IPT-Telemedicine
Monitoring				
	Blood pressure	yes	yes	yes
	Weight	yes	yes	yes
	Heart frequency	yes	yes	yes
	Electrocardiography	no	yes	yes
Questions				
	Symptoms	yes	yes	yes
	Knowledge about heart failure	yes	yes	yes
	Change of behavior	yes	yes	yes
Informing patient about…				
	Symptoms	yes	yes	yes
	Knowledge about heart failure	yes	yes	yes
	Change of behavior	yes	yes	yes
Communication				
	Datacenter	yes	yes	yes
	Medical service center	yes	no	yes
	Direct feedback, true application to patient	yes, through television	yes	yes
	Direct feedback from health care provider to patient	yes, by phone	yes, by phone	yes, by phone
	Continue feedback to health care provider	yes, through software on desktop	yes, through software on desktop	yes, through portal
	Alerts in case of deviation from predefined measurements	yes, through software on desktop	yes, risk profiles (low-middle-high)	yes, through portal
Patient requirements				
	Ability to read	yes	yes	yes
	Active input	yes	yes	yes
	Cognitive functional	yes	yes	yes
	Manual	extensive	simple	simple
	Television	yes	no	no

The 12 clinics that intended to use telemonitoring within a year mostly reported (42%, n=5) that they planned to use the Motiva system (Table 2). The number of patients using telemonitoring in a clinic varied between 10 and 50, but in two clinics more than 75 patients used telemonitoring.

Table 2. Availability and use of telemonitoring (TM) system by actual users (n=31) and planned users (n=12).

TM systems	Actually used system (n=31 clinics)	System of choice in case of a new decision (n=31 clinics)	No current user but expecting to make a choice within 1 year (n=12 clinics)
Health Buddy	7 (28%)	2 (8%)	–
Motiva	14 (46%)	4 (12%)	5 (42%)
IPT Telemedicine	6 (15%)	2 (6%)	–
Other systems	4 (11%)	3 (10%)	2 (16%)
No choice yet	–	4 (12%)	2 (16%)
Unsure	–	16 (52%)	3 (26%)

The following main goals for implementing telemonitoring were reported: “monitoring physical condition”, “monitoring signs of deterioration” (91%, n=39), “monitoring treatment” (74%, n=32), “adjusting medication” (56%, n=24), and “educating patients” (77%, n=33) (see Table 3). Beside these goals, most clinics also used this as a practical reason to start telemonitoring.

Table 3. General descriptive data of heart failure centers using (n=31) and planning to use (n=12) telemonitoring (TM).

Question (n)	Response option	Response n (%)
Number of patients in TM care (n=31 clinics)		
	None	2 (6%)
	0-10	5 (16%)
	10-20	8 (26%)
	20-50	11 (35%)
	50-75	3 (11%)
	>75	2 (6%)
Main goal of using telemonitoring (n=43 clinics, more than one answer possible)		
	Monitoring physical conditioning, signs of deterioration	39 (91%)
	Monitoring and adjustment of treatment	32 (74%)
	Titration of medication	24 (56%)
	Patient education	33 (77%)
	Other goals	3 (7%)
Duration of applying telemonitoring in patient care (n=31 clinics)		
	Between 3 and 6 months	6 (19%)
	Between 6 and 12 months	6 (19%)
	No limit	9 (30%)
	As long as necessary	10 (32%)

Experience With Telemonitoring

Patient Profile

The criteria for using telemonitoring for a specific patient were reported to be based on “needing education” (68 %, n=29), “increasing self management” (63%, n=27), “having complaints of heart failure symptoms” (60%, n=26), and “being (re) admitted due to heart failure” (60%, n=26). See Table 4.

Table 4. Criteria for applying telemonitoring in heart failure (HF) patients.

Criteria for applying telemonitoring	n=43 clinics
Education	29 (68%)
Patient management	27 (63%)
Heart failure re-admission	26 (60%)
Complaints heart failure symptoms	26 (60%)
Based on actual NYHA class	13 (30%)
Medication status	8 (19%)
Different	2 (4%)

Respondents from 8 clinics reported that the current use or amount of medication were reasons for using telemonitoring. The majority of respondents (85%, n=36) stated that the New York Heart Association (NYHA) functional class was not a reason to start telemonitoring (see Table 5).

Table 5. NYHA class in telemonitoring (NYHA: New York Heart Association classification for heart failure), more than one answer possible.

Question (n)	Response option	Response n (%)
Actual NYHA class of patients currently using telemonitoring (n=31)		
	NYHA I	0 (0%)
	NYHA II	19 (61%)
	NYHA III	27 (87%)
	NYHA IV	5 (15%)
Which NYHA class in your patient population is suitable for applying telemonitoring? (n=43)		
	NYHA I	3 (6%)
	NYHA II	14 (32%)
	NYHA III	18 (41%)
	NYHA IV	10 (23%)
Is the NYHA class decisive for applying telemonitoring? (n=43)		
	Yes	6 (15%)
	No	36 (85%)

In order to determine the best course of therapy, heart failure professionals assess the stage of heart failure according to the New York Heart Association (NYHA) functional classification system (see Table 6). This classification system relates symptoms to everyday activities and the patient's quality of life. The NYHA class is not a determined factor for the application of telemonitoring according to the guidelines.

Table 6. NYHA: New York Heart Association classification for heart failure.

Class	Patient symptoms
Class I (Mild)	No limitation of physical activity. Ordinary physical activity does not cause undue fatigue, palpitation, or dyspnea (shortness of breath).
Class II (Mild)	Slight limitation of physical activity. Comfortable at rest, but ordinary physical activity results in fatigue, palpitation, or dyspnea.
Class III (Moderate)	Marked limitation of physical activity. Comfortable at rest, but less than ordinary activity causes fatigue, palpitation, or dyspnea.
Class IV (Severe)	Unable to carry out any physical activity without discomfort. Symptoms of cardiac insufficiency at rest. If any physical activity is undertaken, discomfort is increased.

Nevertheless, patients in NYHA class II and III were most often reported to be enrolled for telemonitoring, whereas no patients in NYHA class I used telemonitoring. In total, 15% of patients in NYHA class IV used telemonitoring.

Length of Time of Telemonitoring

Most respondents stated that they monitor their patients with telemonitoring "as long as needed" or without a time limit. Six clinics noted a maximum time period for using telemonitoring per patient between 3 and 6 months respectively. In response to the question on whether clinics (n=43) could estimate which of the total percentage of all patients in heart failure care were suitable for telemonitoring, the mean percentage was 10%.

Telemonitoring System

Fifteen of the 31 clinics that actually used telemonitoring stated that if a new selection process were to be put in place, they would choose a different system compared to the system they currently used. Sixteen clinics indicated that they were not sure which system they would choose (see Table 2). Of the 31 clinics, 14 reported that they were satisfied with their current telemonitoring system. The other 16 clinics took a neutral stance, and one user reported to be dissatisfied with the telemonitoring equipment.

Expectations Versus Experienced Outcomes

In Figure 1, the expectations of applying telemonitoring are compared with the experienced outcomes after implementation of telemonitoring. The combined 3 groups of aspects of working with telemonitoring (direct patient-related care, telemonitoring system aspects, and organizational aspects) and 10 of the 11 separate items showed that the actual experiences did not meet the prior expectations. The results showed that users had high expectations of the benefits of using telemonitoring, in particular with respect to direct patient-care aspects (mean 7.4).

Expectations of the system-related aspects (mean 6.8) and organizational aspects (mean 6.0) were also high. However, these high expectations of the use of telemonitoring were not reflected in the actual experiences after implementation. The largest difference was found in the group of organizational aspects (reduction of workload score, 5.9 versus 3.5, *P*<.001) and lowering heart failure–related costs, score 5.8 versus 3.2, *P*<.001). The aspect "keeping up with current developments" was the only one in which a reduction was not significant (score, 7.2 versus 6.8, *P*=.15).

Figure 1. Expectations of applying telemonitoring and experienced differences after implementation of telemonitoring.

Organizing and Financing Telemonitoring

A total of 12 clinics (39%) reported to be in a "start-up" period; whereas the other 19 clinics stated that they had fully integrated telemonitoring in their daily care routine. Rules and protocols on the implementation of the system and responsibility for incoming data were available in 70% of the clinics. Protocols on the acceptable length of time between the moment of incoming patient data and the response of the caregiver (response-reaction time) were available in 60% of the clinics. With respect to financing, 54% of telemonitoring systems were financed by health care insurance companies, 13% by project financing, and 7% by the hospital itself or the cardiology department. The other 26% of the clinics did not give insight into their financing of telemonitoring.

Discussion

The most prominent result of our study was that, although the respondents had high perceptions and expectations of working with telemonitoring, these were not positively reflected in the actual experiences.

The trade-offs directly related to the telemonitoring system were most often addressed, but important trade-offs of telemonitoring concerning direct patient care and organizational aspects were only briefly mentioned or not reported at all. A striking finding is that the majority of responding heart failure clinics stated they were considering the use of a different system than the system currently used. Furthermore, aspects of direct patient care (like monitoring and education) were reported as main goals for implementing telemonitoring.

The dominant criteria for using telemonitoring for a specific patient included “education”, “heart failure (re) admission”, and “complaints of heart failure symptoms”. Thirty percent of the respondents mentioned that the actual NYHA class is a criterion for applying telemonitoring, but at the same time only 15% stated that the NYHA class was decisive for applying telemonitoring. In actual practice, the majority of the patients showed to be in NYHA class II and III. Finally, although 1 out of 10 patients was suitable for telemonitoring, the actual number of patients using telemonitoring was limited in general and the duration of the use of telemonitoring unknown. Despite the increased introduction and use of telemonitoring in heart failure, there has been little research regarding user-related aspects of working with telemonitoring. Therefore, it is unknown to what extent expectations, experiences, and possible difficulties in the implementation process of telemonitoring are present in health care providers working with telemonitoring. In this first study to focus specifically on the application of telemonitoring in heart failure clinics, we showed that heart failure clinics have high expectations of patient care, system, and organizational outcomes of working with telemonitoring.

In an earlier study on the expectations of telemonitoring of caregivers in nursing homes, Chang et al [19] reported that respondents expected the benefits of improved efficiency and quality of care, reduction of medical costs, and a reduced workload. However, experiences of telemonitoring were not measured in the study of Chang et al. Although the evidence for the use of telemonitoring in heart failure patients is still growing [5-8], gaps in knowledge about the use of telemonitoring in heart failure remain [3,20,21]. These gaps in knowledge are mainly caused by the absence of data on adequate patient profiling and the overall cost-effectiveness of telemonitoring.

Despite the presence of conflicting evidence on the usefulness of telemonitoring for heart failure and the lack of data regarding the implementation of telemonitoring, the consequences for health care providers, and the logistic processes in daily practice, more than one-third of all heart failure clinics in the Netherlands have implemented this new technology for some of their heart failure patients. This indicates that health care providers have high expectations of working with telemonitoring and are even willing to start working with telemonitoring in the absence of guidelines, protocols, and solid evidence for its usefulness. The use of telemonitoring, however, is still in its infancy, and many clinics are still searching for a way to provide telemonitoring efficiently and effectively. A similar experience was reported with respect to the selection processes for electronic patient records and other technology tools in health care [22-24]. Users were either extremely positive or negative about their system, and this had a “wait-and-see” effect on potential future users. Negative experiences were reflected in the fact that some users were considering looking for a different system than the system currently used. The need for a different system seems to be primarily driven by the practical usage of the system, which falls short of expectations. Our findings indicate that the actual functionalities of the telemonitoring system itself are of great importance to the respondents. Hence, it is questionable if the feeling of overall disappointment is indeed the result of a failing telemonitoring system or is due to a lack of efficient organization around the implementation of telemonitoring systems.

For future success it is very important to create an efficient organization around a system [13]. In the case of telemonitoring, this means that a system should be integrated in a heart failure clinic in which heart failure nurses [11,25] have a coordinating role and have insight in all aspects of patient care (eg, health care professionals involved, situation at home). Within this setting, the heart failure nurse can take appropriate action on the data received from the telemonitoring system [26,27]. Furthermore, additional training is required in which insight and understanding of receiving data, data handling, evaluating expectations, and effect monitoring are vital [28].

Our data showed that in 61% of the heart failure clinics that actually worked with telemonitoring, it was used only in small cohorts with numbers of 10 to 50 patients. Although this concerns only a limited number of patients, it is important to realize that monitoring 50 heart failure patients (next to the treatment of other heart failure patients) might cause a substantial amount of additional work with respect to logistic adjustment, training on using the system, and the development of protocols on data handling, response time, and treatment. We could therefore predict that implementing telemonitoring will not automatically decrease workload.

In this first study on user-related aspects of telemonitoring, we demonstrated that the optimal use of telemonitoring remains a challenge. The main finding of our research is that a substantial difference exists between prior expectations of telemonitoring and the actual use of telemonitoring in daily practice. The focus on, for instance, optimizing medication by using telemonitoring, however, has been shown to be a promising and cost-effective future application [29,30]. While the use of telemonitoring is still in its infancy, it is important to learn from current experiences, even if it currently concerns only a limited number of telemonitoring systems and patients. Ongoing studies such as the IN TOUCH trial [14] in the Netherlands should provide more evidence about cost-effectiveness and the effects of telemonitoring in combination with different types of disease management in heart failure.

A finding that has to be specifically addressed is that most of the respondents indicated that telemonitoring will be applied as long as needed or can even be used indefinitely. This approach should be critically evaluated. First, it might not be the most cost effective in terms of using equipment and staff. Most intervention studies on the use of telemonitoring were short in follow-up, and therefore there are no data available that support the choice for (life) long use of telemonitoring. Second, ethical issues can be raised about whether or not patients would benefit from lifelong monitoring, regardless of the burden on their personal lives. Other notable findings were that 85% of the respondents indicated that the NYHA functional class was not decisive for the application of telemonitoring and that most patients who received telemonitoring were in NYHA functional classes II and III. Although the optimal patient profile for successful use of telemonitoring has not yet been described, it can be expected that specifically patients with severe and more unstable heart failure are suitable for telemonitoring and would benefit in terms of preventing re-admissions. Considering this, it is remarkable that in daily practice telemonitoring is increasingly used for patient education and for optimizing medication in patients with less severe heart failure.

Limitations

For this study, we used a self-developed questionnaire that was not designed to test the feasibility of a telemonitoring system, but rather to examine both the general considerations and reasons for applying telemonitoring in Dutch heart failure clinics, as well as the organizational aspects these systems address. In this study, we did not focus on possible differences in the perception of working with telemonitoring of heart failure nurses and cardiologists, because the main goal of this study was to explore the expectations and experiences of a heart failure team working with telemonitoring. However, one might predict that the comments of the two separate groups would relate to their characteristics. Although we are aware of the limitations of asking about experiences with telemonitoring retrospectively, the design of this study could not correct for this. To account for this limitation, we have focused in the discussion on the learning aspects of the experiences instead of giving clear-cut conclusions.

Conclusion

This representative study (86 of 109 surveyed Dutch heart failure clinics) showed that one- third of heart failure clinics were using or planned to use telemonitoring as part of their care, albeit in a limited number of patients only. Our survey also showed that telemonitoring is not a success story yet. Respondents did not experience a decreased workload while working with telemonitoring, and prior expectations of introducing telemonitoring were not reflected in actual experiences, possibly leading to disappointment. Criteria for both the optimal duration period of using the telemonitoring system and the targeted patient groups were not established, and the choice for a telemonitoring system seemed to be made on the specifications of the system itself, rather than on organizational issues such as protocols or education of staff. All the suppliers of telemonitoring devices observed in this study provide the services of generating and transferring data from a home environment to a health care environment. Telemonitoring is not a “one size fits all” solution. From a patient point of view [9,10] and supported by the recent European Society of Cardiology heart failure guidelines (2012), we conclude that the optimal profile of patients who might benefit from telemonitoring needs to be further explored. Long-term experiences are necessary to discover the most effective use of telemonitoring in terms of reduction of mortality, re-admissions, and improvement of quality of life.

Acknowledgments

We would like to thank the Netherlands Organization for Applied Scientific Research (TNO) for its input in designing the telemonitoring questionnaire (Ing. AJM Rövekamp, Dr. R Mooij); the Trial Coordination Centre (TCC) for supporting, digitalizing, and processing the telemonitoring questionnaire and data; and Dutch heart failure clinics for their cooperation in this telemonitoring survey. This project was funded by the Dutch Ministry of Health, department of Pharmaceutical Affairs and Medical Technology (GMT).

Conflicts of Interest

RB van Dijk is partner of Curit B.V., a Dutch company involved in the development of information and communication technology–guided disease management systems and telemedicine. The other authors declared no conflict of interest.

References

1. Seto E, Leonard KJ, Cafazzo JA, Barnsley J, Masino C, Ross HJ. Perceptions and experiences of heart failure patients and clinicians on the use of mobile phone-based telemonitoring. J Med Internet Res 2012;14(1):e25 [FREE Full text] [doi: 10.2196/jmir.1912] [Medline: 22328237]
2. 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]
3. Bahaadinbeigy K, Yogesan K, Wootton R. Gaps in the systematic reviews of the telemedicine field. J Telemed Telecare 2010;16(7):414-416. [doi: 10.1258/jtt.2010.100505] [Medline: 20841383]

4. Institute of Medicine (US) Committee on Evaluating Clinical Applications of Telemedicine. In: Field MJ, editor. Telemedicine: a guide to assessing telecommunications in health care. Washington, DC: National Academies Press; 1996.
5. Klersy C, De Silvestri A, Gabutti G, Regoli F, Auricchio A. A meta-analysis of remote monitoring of heart failure patients. J Am Coll Cardiol 2009 Oct 27;54(18):1683-1694. [doi: 10.1016/j.jacc.2009.08.017] [Medline: 19850208]
6. Riley JP, Cowie MR. Telemonitoring in heart failure. Heart 2009 Dec;95(23):1964-1968. [doi: 10.1136/hrt.2007.139378] [Medline: 19923337]
7. Inglis SC, Clark RA, McAlister FA, Stewart S, Cleland JG. Which components of heart failure programmes are effective? A systematic review and meta-analysis of the outcomes of structured telephone support or telemonitoring as the primary component of chronic heart failure management in 8323 patients: Abridged Cochrane Review. Eur J Heart Fail 2011 Sep;13(9):1028-1040. [doi: 10.1093/eurjhf/hfr039] [Medline: 21733889]
8. Giamouzis G, Mastrogiannis D, Koutrakis K, Karayannis G, Parisis C, Rountas C, et al. Telemonitoring in chronic heart failure: a systematic review. Cardiol Res Pract 2012;2012:410820 [FREE Full text] [doi: 10.1155/2012/410820] [Medline: 22720184]
9. 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]
10. Chaudhry SI, Mattera JA, Curtis JP, Spertus JA, Herrin J, Lin Z, et al. Telemonitoring in patients with heart failure. N Engl J Med 2010 Dec 9;363(24):2301-2309 [FREE Full text] [doi: 10.1056/NEJMoa1010029] [Medline: 21080835]
11. Desai AS, Stevenson LW. Connecting the circle from home to heart-failure disease management. N Engl J Med 2010 Dec 9;363(24):2364-2367. [doi: 10.1056/NEJMe1011769] [Medline: 21080836]
12. Paré G, Moqadem K, Pineau G, St-Hilaire C. Clinical effects of home telemonitoring in the context of diabetes, asthma, heart failure and hypertension: a systematic review. J Med Internet Res 2010;12(2):e21 [FREE Full text] [doi: 10.2196/jmir.1357] [Medline: 20554500]
13. Joseph V, West RM, Shickle D, Keen J, Clamp S. Key challenges in the development and implementation of telehealth projects. J Telemed Telecare 2011;17(2):71-77. [doi: 10.1258/jtt.2010.100315] [Medline: 21097563]
14. de Vries AE, de Jong RM, van der Wal MH, Jaarsma T, van Dijk RB, Hillege HL. The value of INnovative ICT guided disease management combined with Telemonitoring in OUtpatient clinics for Chronic Heart failure patients. Design and methodology of the IN TOUCH study: a multicenter randomised trial. BMC Health Serv Res 2011;11:167 [FREE Full text] [doi: 10.1186/1472-6963-11-167] [Medline: 21752280]
15. Philips. Motiva URL: http://www.healthcare.philips.com/main/products/telehealth/products/motiva.wpd [accessed 2012-08-28] [WebCite Cache ID 6AFsEVR3F]
16. Sananet. Health Buddy URL: http://www.sananet.nl/index.php?url=/technologie/het_health_buddy_systeem/ [accessed 2012-08-28] [WebCite Cache ID 6AFrtRlb7]
17. IPT Telemedicine. URL: http://www.ipt-telemedicine.nl/ [accessed 2012-08-28] [WebCite Cache ID 6AFsS6dxJ]
18. van Leven K, Rövekamp T, Valentijn P. Inventarisatie eHealthNu Expertgroep Hartfalen. The Netherlands: eHealthNu; 2010.
19. Chang JY, Chen LK, Chang CC. Perspectives and expectations for telemedicine opportunities from families of nursing home residents and caregivers in nursing homes. Int J Med Inform 2009 Jul;78(7):494-502. [doi: 10.1016/j.ijmedinf.2009.02.009] [Medline: 19345640]
20. Koehler F, Winkler S, Schieber M, Sechtem U, Stangl K, Böhm M, et al. Telemedicine in heart failure: Pre-specified and exploratory subgroup analyses from the TIM-HF trial. Int J Cardiol 2012 Nov 29;161(3):143-150. [doi: 10.1016/j.ijcard.2011.09.007] [Medline: 21982700]
21. Angermann CE, Störk S, Gelbrich G, Faller H, Jahns R, Frantz S, Competence Network Heart Failure. Mode of action and effects of standardized collaborative disease management on mortality and morbidity in patients with systolic heart failure: the Interdisciplinary Network for Heart Failure (INH) study. Circ Heart Fail 2012 Jan;5(1):25-35. [doi: 10.1161/CIRCHEARTFAILURE.111.962969] [Medline: 21956192]
22. Lorenzi NM, Riley RT. Organizational impact of health information systems in healthcare. Stud Health Technol Inform 2002;65:396-406. [Medline: 15460238]
23. Lorenzi NM, Riley RT. Knowledge and change in health care organizations. Stud Health Technol Inform 2000;76:63-69. [Medline: 10947502]
24. Colpas P. Digital dilemma. Experts provide tips on purchasing electronic health records systems. Health Manag Technol 2010 Jun;31(6):12-14. [Medline: 20593695]
25. McDonagh TA, Blue L, Clark AL, Dahlström U, Ekman I, Lainscak M, European Society of Cardiology Heart Failure Association Committee on Patient Care. European Society of Cardiology Heart Failure Association Standards for delivering heart failure care. Eur J Heart Fail 2011 Mar;13(3):235-241. [doi: 10.1093/eurjhf/hfq221] [Medline: 21159794]
26. Lorenzi NM, Riley RT. Managing change: an overview. J Am Med Inform Assoc 2000;7(2):116-124 [FREE Full text] [Medline: 10730594]
27. Lorenzi NM, Riley RT. Organizational issues = change. Int J Med Inform 2003 Mar;69(2-3):197-203. [Medline: 12810124]

28. Waldman JD, Yourstone SA, Smith HL. Learning curves in health care. Health Care Manage Rev 2003;28(1):41-54. [Medline: 12638372]
29. de Vries AE, van der Wal MH, Bedijn W, de Jong RM, van Dijk RB, Hillege HL, et al. Follow-up and treatment of an instable patient with heart failure using telemonitoring and a computerised disease management system: a case report. Eur J Cardiovasc Nurs 2012 Dec;11(4):432-438. [doi: 10.1016/j.ejcnurse.2011.04.001] [Medline: 21546311]
30. Antonicelli R, Mazzanti I, Abbatecola AM, Parati G. Impact of home patient telemonitoring on use of β-blockers in congestive heart failure. Drugs Aging 2010 Oct 1;27(10):801-805. [doi: 10.2165/11538210-000000000-00000] [Medline: 20883060]

Edited by G Eysenbach; submitted 17.05.12; peer-reviewed by L Grieten, E Gabarron; comments to author 27.06.12; revised version received 27.07.12; accepted 24.09.12; published 10.01.13.

Please cite as:
de Vries AE, van der Wal MH, Nieuwenhuis MM, de Jong RM, van Dijk RB, Jaarsma T, Hillege HL
Health Professionals' Expectations Versus Experiences of Internet-Based Telemonitoring: Survey Among Heart Failure Clinics
J Med Internet Res 2013;15(1):e4
URL: http://www.jmir.org/2013/1/e4/
doi:10.2196/jmir.2161
PMID:23305645

Original Paper

Challenges for Researchers Investigating Contraceptive Use and Pregnancy Intentions of Young Women Living in Urban and Rural Areas of Australia: Face-to-Face Discussions to Increase Participation in a Web-Based Survey

Danielle L Herbert[1], PhD; Deborah Loxton[2], PhD; Deborah Bateson[3], MBBS; Edith Weisberg[3], MBBS, FRANZCOG; Jayne C Lucke[1], PhD

[1]UQ Centre for Clinical Research, University of Queensland, Brisbane, Australia
[2]Research Centre for Gender, Health and Ageing, University of Newcastle, Newcastle, Australia
[3]Family Planning NSW, Sydney, Australia

Corresponding Author:
Jayne C Lucke, PhD
UQ Centre for Clinical Research
University of Queensland
Building 71/918, Level 7
Royal Brisbane & Women's Hospital Site, Herston Road,
Brisbane, 4029
Australia
Phone: 61 733466031
Fax: 61 733465598
Email: j.lucke@uq.edu.au

Abstract

Background: It is imperative to understand how to engage young women in research about issues that are important to them. There is limited reliable data on how young women access contraception in Australia especially in rural areas where services may be less available.

Objective: This paper identifies the challenges involved in engaging young Australian women aged 18-23 years to participate in a web-based survey on contraception and pregnancy and ensure their ongoing commitment to follow-up web-based surveys.

Methods: A group of young women, aged 18-23 years and living in urban and rural New South Wales, Australia, were recruited to participate in face-to-face discussions using several methods of recruitment: direct contact (face-to-face, telephone or email) and snowball sampling by potential participants inviting their friends. All discussions were transcribed verbatim and analyzed using thematic analysis.

Results: Twenty young women participated (urban, n=10: mean age 21.6 years; rural, n=10: 20.0 years) and all used computers or smart phones to access the internet on a daily basis. All participants were concerned about the cost of internet access and utilized free access to social media on their mobile phones. Their willingness to participate in a web-based survey was dependent on incentives with a preference for small financial rewards. Most participants were concerned about their personal details and survey responses remaining confidential and secure. The most appropriate survey would take up to 15 minutes to complete, be a mix of short and long questions and eye-catching with bright colours. Questions on the sensitive topics of sexual activity, contraception and pregnancy were acceptable if they could respond with “I prefer not to answer”.

Conclusions: There are demographic, participation and survey design challenges in engaging young women in a web-based survey. Based on our findings, future research efforts are needed to understand the full extent of the role social media and incentives play in the decision of young women to participate in web-based research.

(J Med Internet Res 2013;15(1):e10) doi:10.2196/jmir.2266

KEYWORDS

Rural; urban; Australia; contraception; pregnancy; participant recruitment; women's health; Facebook; social media; web-based survey

Introduction

Little information is available about how young women use and access contraception in any region of Australia, urban or rural. The need for information in this area is long overdue, as has been noted in the literature [1,2]. Existing Australian research into sexual activity, contraception, and pregnancy includes the National Survey of Australian Secondary Students from 1992 to 2008, prior to the students leaving school [3], and the Australian Study of Health and Relationships that surveyed adults aged 16-64 years [4,5]. Further research includes the Australian Longitudinal Study on Women's Health (ALSWH) of the cohort born in 1973-78. They have been followed since 1996 at age 18-23 years [6]. Additional maternal health data from the birth of the first child to the end of women's reproductive lives are reported annually in Australia [7]. However, there remains limited research specifically on the contraceptive use and pregnancy intentions of the new generation of young women.

It is difficult to capture the voice of young people living in rural areas in large population studies. For example, in 2010, New South Wales (NSW, the most populous state in Australia) had a population of 7.2 million, of which 14% were aged 15-24 years [8]. An analysis of population data highlighted demographic challenges as most young women in NSW live in major cities: only 16% of women aged 15-24 years lived in inner regional areas and 5% lived in outer regional and remote areas [8,9]. From age 20 years, the proportion of young women living outside the major metropolitan area of Sydney decreased across all other areas of NSW, indicating the movement of young women from rural areas to the greater Sydney region. Age-specific fertility rates in NSW are lower than the national rate in Australia, respectively, for 15-19 years (12.9 and 15.5 births per 1000 women) and 20-24 years (49.3 and 52.5 births per 1000 women) [7,8]. However, the fertility rates are dramatically higher in rural areas compared to rates in urban areas. While web-based recruitment can overcome any geographical limitations, response rates to population studies cannot be calculated using established methods because the number of potential participants is unknown [10].

There are many challenges in researching the health of young people, especially those living in rural areas [11]. Galea and Tracy (2007) have reported on the reasons for declining participation in epidemiologic studies and the resulting difficulty in recruiting population-representative cohorts: people are less likely to participate if they do not see the direct relevance to their personal lives. However, women who are employed and have a higher socioeconomic status are more likely to participate [12]. Differences in response rates in population-based studies may be accounted for by the sampling strategy used to identify eligible persons [12,13], including web-based recruitment through Facebook advertising [14]. A modern challenge for engaging young people in research is their preference for social media and the pursuit of a digital life [15]. An Australian survey of household Internet subscribers in 2010-2011 found social networking was one of the most popular Internet activities [16]. Almost all (96%) persons aged 18-24 years had access to the Internet and used their connection for emailing (93%) or social networking and online gaming (86%) [16]. A 2009 survey of 196,000 young people aged 15-24 years and living with a disability found 86% had access to the Internet at home [16]. Studies of the Internet activity of Australian households show that almost all households access the Internet either daily or weekly [16]. The use of social media by young Australian adults is similar to young American adults [17].

Little is known about how young women access contraception, particularly in rural areas of Australia. It is assumed that many young women visit their local GP for advice about contraception. In Australia, from 1998, standard general practice consultations have been monitored through the Bettering the Evaluation and Care of Health (BEACH) program [18]. Taylor et al (2010) analyzed BEACH data from 1994 to 2009 and found a reversal in the historical trend for extended consultations and the emerging preference for short appointments suited to single-issue consultations [19]. However, the comprehensive assessment of an individual's need for contraception and pregnancy advice may require extended multi-issue consultations. Although a prescription to continue the oral contraceptive pill requires a short consultation, women who need to change their contraceptive method and be informed about new methods are likely to require an extended consultation or multiple visits.

Taylor et al (2010) recognized that a limitation of their study was the focus on individual services and not the assessment of individuals and their longitudinal use of services [19]. In 2001, adolescents aged 15-24 years had fewer visits to a doctor compared to older Australians, but most young women were attending for issues related to contraception or pregnancy [20]. Mazza et al (2012) used BEACH data to analyze contraceptive management by GPs and found a rate of 60 per 1000 consultations focused on contraception for women aged 12-54 years regardless of their residence in urban or rural Australia [21]. The extent to which young women get contraceptive information, advice, and services from locations other than their local GP is not known.

In order for research findings about how young women use and access contraception to inform policy and practice, it is important to engage young women in research design and implementation. This study aims to identify the challenges involved in engaging young women aged 18-23 years, especially those from rural areas, to participate in a web-based survey on contraception and pregnancy. Further, this study investigates the challenges involved in engaging young women to complete multiple follow-up web-based surveys in order to study changes over time in contraceptive use and pregnancy intentions.

Methods

Study Design

This study is part of a larger project called Contraceptive Use, Pregnancy Intention and Decisions (CUPID) of young Australian women. CUPID aims to shed light on the reasons for using, or not needing, contraceptive methods among young women, despite the apparent widespread availability of contraception. This study aims to identify how to encourage the participation of young women in a web-based survey, about contraception and pregnancy, through face-to-face discussions with young women. Ethical approvals were granted by committees at the University of Queensland, University of Newcastle, and Family Planning NSW. Additional information on CUPID is available from the project website.

Sample and Recruitment

Young women aged 18-23 years and living in urban and rural NSW, Australia, were invited to participate in focus group discussions. Advertising for the discussions was via posters and information sheets displayed at the University of Newcastle, Family Planning NSW clinics (Ashfield, Fairfield, Penrith, Newcastle, Dubbo), educational settings (library), and one workplace setting with a trainee program including young women aged 18-23 years. Recruitment was achieved using direct contact (face-to-face, telephone, or email) and potential participants inviting their friends (ie, snowballing sampling). No financial incentives were given for participation in the study, but refreshments were provided during the discussions. Women who offered to participate but were aged outside the specified age range were excluded.

Procedure

Ten semi-structured focus group discussions (lasting between 30 and 60 minutes) were conducted with participants from August to September 2011 by one member of the research team (DH) and an assistant (CR). The focus groups were held in urban areas at clinical locations in Sydney (Ashfield, Penrith) and Newcastle. In rural areas, the focus groups were held at clinical (Dubbo), educational (Tamworth), and workplace (Muswellbrook) locations local to the participants. Interviews were conducted with single participants where it was not possible to form a focus group of two or more participants aged 18-23 years, or the participant specifically requested to have a one-on-one discussion. The interview followed the focus group question schedule.

Before commencing the discussion, participants were given an information sheet describing the aims of the study, a consent form, and short survey to gather information on age, study or work patterns, and Internet use. All discussions were audio recorded after written consent was provided by the participants. A question schedule was used to stimulate discussion and included open-ended questions to encourage participants to give a broad range of responses. The interviewer (equally DH or CR) prompted the participants where appropriate to elicit more detailed information. Participants were asked to avoid revealing their personal experience of contraception and pregnancy but to focus on their reaction to survey questions on these topics. Other topics included access to the Internet, preference for using computers, smartphones or other devices, and responding to sensitive questions. The non-interviewing researcher (equally DH or CR) took notes throughout and summarized the discussion to the participants at the end of each discussion and invited corrections to the recording.

Coding and Analysis

All discussions were de-identified, transcribed verbatim, and manually analyzed using thematic analysis. Our overarching aim was to identify important recommendations for research about contraception and unplanned pregnancy. We were interested in examining broad themes with which to frame recommendations, rather than detailed content analysis. We therefore employed a grounded theory approach [22] to examine the comments provided in the interview about participation challenges and survey challenges. Printed interview transcripts were read by DH. Statements that, directly or indirectly, referred to participation or survey challenges were identified. Excerpts were sorted into initial groupings by DH. These excerpts revealed five themes relating to participation challenges and seven themes relating to survey challenges. The research team reviewed these themes and agreed on the coding framework. All interviews were then coded by DH according to the identified themes using NVivo 9 (QSR International Pty Ltd., 2010). JL reviewed the transcripts and coding in detail and any minor discrepancies were resolved, and interpretations developed, with in-depth discussion to reach a consensus.

Results

Twenty young women aged 18-23 years participated in the discussions (Table 1). Focus groups were conducted with five groups of participants at Dubbo (2 participants in a clinical setting), Newcastle (2 participants in a clinical setting), Sydney (one group of 2 participants and one group of 3 participants, both in a clinical setting,) and Muswellbrook (6 participants in a workplace setting). Due to recruitment difficulties or participant request, interviews were conducted with single participants at Tamworth (2 participants in an educational setting), Sydney (1 participant in a clinical setting), and Newcastle (2 participants in a clinical setting). Participants were living in urban (n=10) and rural (n=10) NSW; mean age 20.8 years. Half (n=10) of the participants were working only and not continuing with study after high school: most of these young women lived in rural areas (7/10). Participants from rural NSW were younger, working rather than studying, and unlikely to be attending university compared to participants from urban NSW. All participants used either computers or smartphones to access the Internet on a daily basis.

Two key domains were identified from the discussions: participation challenges to encourage young women to participate in a web-based survey and survey design challenges to ensure they would complete a survey including sensitive questions about sexual activity, contraceptive use, and pregnancy (Table 2). Direct quotes are included to provide contextual information for the identified domains.

Table 1. Characteristics of participants living in New South Wales (NSW, Australia).

Categories		Urban NSW n=10	Rural NSW n=10	Total n=20 (%)
Age				
	18-19	0	5	5 (25)
	20-21	6	2	8 (40)
	22-23	4	3	7 (35)
Occupation				
	Studying	6	1	7 (35)
	Working	3	7	10 (50)
	Studying & working	1	0	1 (5)
	Full-time mother	0	2	2 (10)
Highest qualification				
	Middle high school	0	2	2 (10)
	Senior high school	4	5	9 (45)
	Technical college	0	2	2 (10)
	University	6	1	7 (35)
Internet use				
	Daily	10	10	20 (100)
Internet access				
	Computer	3	4	7 (35)
	Smartphone	0	2	2 (10)
	Computer & smartphone	7	4	11 (55)

Table 2. Thematic analysis of focus group discussions (General = 19-20 participants, Typical = 10-18 participants, Variant = 1-9 participants).

Domains and subdomains		Themes	Frequency
Participation challenges			
	Internet access	Connect to Internet surveys	General
		Low coverage in some rural areas	
		Cost of access and download quota	
	Social media	Daily use of Facebook instead of email	General
		Free access to Facebook on mobiles	
	Incentives	Prefer financial incentive	General
		Want to know chance of winning prizes	
	Privacy	Confidentiality, trust, and security	Typical
		Reluctance to provide contact details	
	Prior knowledge	Want to know about the study before being asked to participate	Variant
Survey challenges			
	Format	Design of questions	Typical
		Mix of short and long questions	
		Up to 15 minutes to complete survey	
		Eye-catching and bright colors	
	Content	Sensitive questions okay if given option “prefer not to answer”	Typical
	Relationships	Range of potential relationship situations	Variant
		Committed versus casual relationships	
	Contraception	Some women use the term “protection”	Variant
		Never heard of some contraceptive methods	
	Pregnancy	Relationship status associated with pregnancy intentions	Variant
	Feedback	Want summary information of findings	Variant
		Like to compare themselves to the summary	
		Use of summary to inform their local area	
	Reminders	Continually receiving SMS reminders from range of companies	Variant
		Will tolerate 2-3 reminders to complete follow-up surveys but restrict to 1/week	
		Need a deadline to prompt survey completion	

Participation Challenges

A recurring theme in the discussions was the ongoing cost of using the Internet, either on a smartphone, mobile phone or home computer: “But yes, being on the Internet and it’s going to cost you money ... because the Internet is pricey ... it’s just once you’re on there it’s like tick, tick, tick, money gone.” The constant awareness of the increasing Internet charges discouraged participation in a web-based survey. However, free access to social media through most mobile network service providers resulted in most participants using social media for almost all personal communications: “At the moment because I’m with [mobile network provider] ... you get free Facebook access, so that’s the only reason why I go on Facebook, because it’s free.” The majority of participants suggested the use of social media as the “only way” to engage their attention and raise awareness of the web-based survey.

An incentive to participate was desired by almost all participants: “I think it’d probably have to be like $20 [gift card] ... I do think that people do sometimes get a bit greedy, and it’s pretty sad that people just won’t do it to help people out.” Privacy and confidentiality were key concerns about participating in web-based surveys. Participants were concerned about providing their contact details especially if there would be future unexpected contact: “... and they want to interview me or, I don’t know, talk to me over the phone, then I wouldn’t necessarily give it.” Further, participants were worried about who wanted their information: “It’s not a freaky guy behind the computer just trying to get information about you”. Participants also wondered how their responses would be used: “Just you’re worried about where your information is going to go, and who’s going to get it.”

Survey Design Challenges

Participants wanted the survey to be as short as possible (15 minute maximum) otherwise they were likely to leave the web-based survey and move to other websites. Asking sensitive questions was appropriate as long as "I prefer not to answer" was given as a response option:

> *You might still get people who are going to deny it when it's true, because I've probably lied in questionnaires before about - oh sorry - but certain questions I was like, well, I don't want to answer that. People would rather put a lie than leave it blank.*

In the absence of this response option, most participants admitted they would give a false response in the survey.

Participants wanted to receive a combined feedback and reminder to complete the follow-up surveys because this information indicated their prior participation had contributed to some results: "… if you get some results back so you know... this is what your information gave, so then the next time they know it's actually going to do something ... they're more willing to do the survey in the next six months." Participants from rural areas identified the benefit of using the feedback to advocate for change in their own community:

> *Or you did show a comparison because it's rural and city. Show us the comparison. We're rural, we want to see where we're lacking and why we haven't got the same things so that maybe, I don't know how, but we could pitch it towards our community to try and get it going.*

Many participants suggested certain personality types who were proactive, organized, and appreciated the benefits of research would always be more likely to participate in a web-based survey than other types of young women.

Discussion

Principal Findings

This study examined two key challenges to researchers in engaging young Australian women in studies of contraception and pregnancy. These areas are (1) ensuring adequate participation and (2) appropriate survey design to ensure the web-based survey is completed. Challenges relating to achieving adequate participation in a survey included the adequacy of participants' access to the Internet, the use of social media, and avoidance of additional financial costs to access the web-based survey for young women living with socioeconomic disadvantages. Financial incentives were identified as the most effective method of ensuring recruitment of participants for short web-based surveys. Survey design challenges included the need for the survey to be short, to include the option to not answer questions, and for participants to receive feedback about the results. These challenges indicated that web-based surveys should be tailored to the needs of young women, especially those living in lower socioeconomic circumstances.

Implications for Researchers

The web-based recruitment of research participants have been found to overcome the reluctance of young people to attend face-to-face discussions [23] and can reach hidden populations, eg, people with disabilities [10]. The use of the most popular social networking site, Facebook, has been identified as effective for web-based recruitment [14]. Fenner et al (2012) found that from 7940 Australians who clicked on the Facebook advertisement of their project, 3.5% went on to participate in the study. Among the young female participants, 28% were from a rural area, which was higher than the 21% of the target population from a rural area [14]. A US study of adult substance use also recruited through targeted Facebook advertising and successfully recruited 10% (n=1548) of the 14,808 people who clicked on the Facebook advertising, met the inclusion criteria and completed the survey [17]. However, web-based recruitment cannot occur when there is a digital divide beyond economics, eg, people living in rural Australia with unreliable or non-existent Internet availability. Notwithstanding these limitations, the promotion of research participation to young people as a means to have a voice, be heard, and contribute to their community may be the key driver to participate [24]. The findings of this study have shown that young women, especially those in rural areas, may be willing to participate in web-based research on issues important to them that will help them in their local community. Further, the clear message from young women was the need for a social media presence to grab their attention and ensure their participation.

One option that has been used to counteract the demographic challenge presented by the unequal numbers of young women in rural and urban areas of Australia is the use of a stratified random sampling frame to ensure young women living in rural areas are well represented and the findings are not urban-centric. Such a sampling frame is an established method, eg, the ALSWH has a successful history of recruiting and maintaining three generational cohorts of Australian women using mailed invitations and surveys [6,25]. However, the preference of young women to use social media or email instead of postal mail limits the effectiveness of using a similar methodology in a new generation of young women [14,15]. Participants in the ALSWH and in the Fenner et al (2012) study were broadly representative of the Australian female population with recognized over-representation of higher educated women whether the participants responded to postal invitations [6] or Facebook advertising [14]. A limitation of web-based recruitment, eg, social media, and subsequent volunteer participation is the potential loss of sample representativeness and generalizability of the findings. However, a very low response to postal invitations based on a stratified random sampling frame would also result in a non-representative sample of young women. A challenge for researchers is to adapt to generational changes within their target population and choose the least limiting web-based methodology while ensuring sufficient statistical power from their sample.

This study has identified key design features of a web-based survey to increase the likelihood of young women's ongoing commitment to follow-up web-based surveys. Young women need to be reassured that their anonymity and confidentiality are maintained. They need to feel supported and secure that the researchers conducting the survey are legitimate and trustworthy by having recognizable logos and contact details. Young women agreed that university and government logos were seen as more legitimate than other novel logos viewed on websites, and these logos would encourage their participation. Feedback was imperative to ensure their continued commitment to the research. A study from the University of Wisconsin, USA, showed the suggestion of sending future email reminders to finish a partially completed survey was sufficient to encourage participants to immediately finalize the survey [26]. Further, the findings reported from this current study showed young women had a limited tolerance of reminders and preferred a short survey to be completed as quickly as possible.

This study successfully engaged a group of young women from rural and urban NSW in the development of a web-based survey of contraception and unplanned pregnancy. The findings of these discussions have implications for the research design and implementation of web-based surveys about sexual activity, contraception, and pregnancy, which must include potentially sensitive questions [27]. It is imperative to offer young women an appropriate range of response options to avoid forced responses that create false or blank reporting [27]. Reliable data are needed in order to make recommendations for policy and clinical practice interventions to reduce rates of unintended pregnancy.

Conclusions

Information about contraception use and pregnancy intention in young Australian women is essential to inform policy but data are lacking, especially for those living in rural areas. Based on our findings, future research efforts are needed to understand the full extent of the role social media and incentives play in the decision of young women to participate in web-based research.

Acknowledgments

We gratefully acknowledge funding support from the Australian Research Council (Linkage Project Grant LP100200349), and partner organizations Family Planning NSW and Bayer Australia Ltd. The research on which this paper is based was conducted at the Universities of Queensland and Newcastle. We are grateful to the women who participated in the discussions. We would also like to thank Clare Rooney (CR) for her help with conducting the focus groups.

Conflicts of Interest

The research is funded by an ARC Linkage Grant that involves some cash and in-kind support from Family Planning NSW and Bayer Australia Ltd. The lead author (DH) was employed through the Linkage Project Grant. Deborah Loxton, Edith Weisberg, and Jayne Lucke are Investigators on the grant. Two co-authors (DB and EW) are employees of Family Planning NSW.

References

1. Taft AJ, Hobbs MK, Hussainy SY, Amir LH, Stewart K, Smith AM, et al. Unintended pregnancy in Australia: what more can we do? Med J Aust 2011 Aug 15;195(4):166-167. [Medline: 21843114]
2. Lucke JC, Herbert DL, Loxton D, Weisberg E. Unintended pregnancies: reducing rates by improving access to contraception. Aust Fam Physician 2011 Nov;40(11):849 [FREE Full text] [Medline: 22059210]
3. Smith A, Agius P, Mitchell A, Barrett C, Pitts M. La Trobe University. Melbourne, Australia; 2009. Secondary Students and Sexual Health 2008: Results of the 4th National Survey of Australian Secondary Students, HIV/AIDS and Sexual Health URL: http://www.latrobe.edu.au/arcshs/downloads/arcshs-research-publications/secondary-students-and-sexual-health-2008.pdf [accessed 2013-01-07] [WebCite Cache ID 6DV4AerfE]
4. Smith AM, Rissel CE, Richters J, Grulich AE, de Visser RO. Sex in Australia: the rationale and methods of the Australian Study of Health and Relationships. Aust N Z J Public Health 2003;27(2):106-117. [Medline: 14696700]
5. Smith AM, Pitts MK, Shelley JM, Richters J, Ferris J. The Australian longitudinal study of health and relationships. BMC Public Health 2007;7:139 [FREE Full text] [doi: 10.1186/1471-2458-7-139] [Medline: 17608953]
6. Lee C, Dobson AJ, Brown WJ, Bryson L, Byles J, Warner-Smith P, et al. Cohort Profile: the Australian Longitudinal Study on Women's Health. Int J Epidemiol 2005 Oct;34(5):987-991 [FREE Full text] [doi: 10.1093/ije/dyi098] [Medline: 15894591]
7. Australian Bureau of Statistics (ABS). 3301.0 - Births Australia 2010. Canberra, Australia; 2011 Oct 25. URL: http://www.abs.gov.au/ausstats/abs@.nsf/mf/3301.0/ [accessed 2012-07-11] [WebCite Cache ID 695ik9LhL]
8. Australian Bureau of Statistics (ABS). National Regional Profile: New South Wales at 30 June 2010. Canberra, Australia; 2011 Nov 04. URL: http://www.ausstats.abs.gov.au/ausstats/nrpmaps.nsf/NEW%20GmapPages/national%20regional%20profile [accessed 2012-07-11] [WebCite Cache ID 695jMnwky]
9. Australian Bureau of Statistics (ABS). 1244.0.00.001 - Information Paper : Outcomes of ABS Views on Remoteness Consultation, Australia, Jun 2001. Canberra, Australia; 2001 Jul 05. URL: http://www.abs.gov.au/ausstats/abs@.nsf/mf/1244.0.00.001 [accessed 2012-07-11] [WebCite Cache ID 695jWtD32]
10. 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]
11. Jeffery V, Ervin K. Responding to rural health needs through community participation: addressing the concerns of children and young adults. Aust J Prim Health 2011;17(2):125-130. [doi: 10.1071/PY10050] [Medline: 21645466]

12. Galea S, Tracy M. Participation rates in epidemiologic studies. Ann Epidemiol 2007 Sep;17(9):643-653. [doi: 10.1016/j.annepidem.2007.03.013] [Medline: 17553702]
13. Morton SM, Bandara DK, Robinson EM, Carr PE. In the 21st Century, what is an acceptable response rate? Aust N Z J Public Health 2012 Apr;36(2):106-108. [doi: 10.1111/j.1753-6405.2012.00854.x] [Medline: 22487341]
14. 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]
15. Watkins SC. The young and the digital: what the migration to social-network sites, games, and anytime, anywhere media means for our future. Boston: Beacon Press; 2009.
16. Australian Bureau of Statistics (ABS). 8146.0 - Household Use of Information Technology, Australia, 2010-11. Canberra, Australia; 2011 Dec 15. URL: http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/8146.0Main%20Features12010-11?OpenDocument [accessed 2012-10-09] [WebCite Cache ID 6BHF60LTr]
17. 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]
18. Britt H, Miller G, Charles J, Henderson J, Bayram C, Pan Y, et al. General practice activity in Australia 2010-11. BEACH: Bettering the Evaluation and Care of Health. Sydney, Australia: The University of Sydney; 2011 Nov. General Practice Series Number 29 URL: http://ses.library.usyd.edu.au/bitstream/2123/7772/4/9781920899868_CDROM.pdf [accessed 2012-07-11] [WebCite Cache ID 695iwGjcd]
19. Taylor MJ, Horey D, Livingstone C, Swerissen H. Decline with a capital D: long-term changes in general practice consultation patterns across Australia. Med J Aust 2010 Jul 19;193(2):80-83. [Medline: 20642408]
20. Booth ML, Knox S, Kang M. Encounters between adolescents and general practice in Australia. J Paediatr Child Health 2008 Dec;44(12):699-705. [doi: 10.1111/j.1440-1754.2008.01409.x] [Medline: 19077068]
21. Mazza D, Harrison C, Taft A, Brijnath B, Britt H, Hobbs M, et al. Current contraceptive management in Australian general practice: an analysis of BEACH data. Med J Aust 2012 Jul 16;197(2):110-114. [Medline: 22794058]
22. Patton MQ. Variety in qualitative inquiry: theoretical orientations. In: Qualitative research and evaluation methods (3rd edition). Thousand Oaks, California: Sage Publications; 2002:75-142.
23. Fox FE, Morris M, Rumsey N. Doing synchronous online focus groups with young people: methodological reflections. Qual Health Res 2007 Apr;17(4):539-547. [doi: 10.1177/1049732306298754] [Medline: 17416707]
24. Rheingold H. Using participatory media and public voice to encourage civic engagement. In: Bennett WL, editor. Civic life online: learning how digital media can engage youth. Cambridge, MA: The MIT Press; 2008:97-118.
25. Young AF, Powers JR, Bell SL. Attrition in longitudinal studies: who do you lose? Aust N Z J Public Health 2006 Aug;30(4):353-361. [Medline: 16956166]
26. Klofstad CA, Boulianne S, Basson D. Matching the message to the medium: results from an experiment on internet survey email contacts. Soc Sci Comp Rev 2008;26(4):498-509 [FREE Full text]
27. Tourangeau R, Yan T. Sensitive questions in surveys. Psychol Bull 2007 Sep;133(5):859-883. [doi: 10.1037/0033-2909.133.5.859] [Medline: 17723033]

Abbreviations

ALSWH: Australian Longitudinal Study on Women’s Health
BEACH: Bettering the Evaluation and Care of Health
CUPID: Contraceptive Use, Pregnancy Intention and Decisions of young Australian women
NSW: New South Wales, Australia

Edited by G Eysenbach; submitted 12.07.12; peer-reviewed by J Burns, A Taft, E Amedonu, J Dickerson; comments to author 31.08.12; revised version received 09.10.12; accepted 25.10.12; published 21.01.13.

Please cite as:
Herbert DL, Loxton D, Bateson D, Weisberg E, Lucke JC
Challenges for Researchers Investigating Contraceptive Use and Pregnancy Intentions of Young Women Living in Urban and Rural Areas of Australia: Face-to-Face Discussions to Increase Participation in a Web-Based Survey
J Med Internet Res 2013;15(1):e10
URL: http://www.jmir.org/2013/1/e10/
doi:10.2196/jmir.2266
PMID:23337208

Original Paper

Human Computation as a New Method for Evidence-Based Knowledge Transfer in Web-Based Guideline Development Groups: Proof of Concept Randomized Controlled Trial

Annemie Heselmans[1], PhD; Bert Aertgeerts[2,3], MD, PhD; Peter Donceel[4], MD, PhD; Stijn Van de Velde[3], PhD; Peter Vanbrabant[5], MD; Dirk Ramaekers[1,3,6], MD, PhD

[1]School of Public Health and Primary Care, Katholieke Universiteit Leuven, Leuven, Belgium
[2]School of Public Health and Primary Care, Academic Center for General Practice, Katholieke Universiteit Leuven, Leuven, Belgium
[3]Belgian Center for Evidence-Based Medicine, Belgian Branch of the Cochrane Collaboration, Leuven, Belgium
[4]School of Public Health and Primary Care, Occupational, Environmental and Insurance Medicine, Katholieke Universiteit Leuven, Leuven, Belgium
[5]University Hospital Leuven, Emergency department, Katholieke Universiteit Leuven, Leuven, Belgium
[6]ZNA Hospital Network Antwerp, Antwerpen, Belgium

Corresponding Author:
Annemie Heselmans, PhD
School of Public Health and Primary Care
Katholieke Universiteit Leuven
Kapucijnenvoer 35 blok d, 3000
Leuven,
Belgium
Phone: 32 0471654129
Fax: 32 16337480
Email: annemie.heselmans@med.kuleuven.be

Abstract

Background: Guideline developers use different consensus methods to develop evidence-based clinical practice guidelines. Previous research suggests that existing guideline development techniques are subject to methodological problems and are logistically demanding. Guideline developers welcome new methods that facilitate a methodologically sound decision-making process. Systems that aggregate knowledge while participants play a game are one class of human computation applications. Researchers have already proven that these games with a purpose are effective in building common sense knowledge databases.

Objective: We aimed to evaluate the feasibility of a new consensus method based on human computation techniques compared to an informal face-to-face consensus method.

Methods: We set up a randomized design to study 2 different methods for guideline development within a group of advanced students completing a master of nursing and obstetrics. Students who participated in the trial were enrolled in an evidence-based health care course. We compared the Web-based method of human-based computation (HC) with an informal face-to-face consensus method (IC). We used 4 clinical scenarios of lower back pain as the subject of the consensus process. These scenarios concerned the following topics: (1) medical imaging, (2) therapeutic options, (3) drugs use, and (4) sick leave. Outcomes were expressed as the amount of group (dis)agreement and the concordance of answers with clinical evidence. We estimated within-group and between-group effect sizes by calculating Cohen’s d. We calculated within-group effect sizes as the absolute difference between the outcome value at round 3 and the baseline outcome value, divided by the pooled standard deviation. We calculated between-group effect sizes as the absolute difference between the mean change in outcome value across rounds in HC and the mean change in outcome value across rounds in IC, divided by the pooled standard deviation. We analyzed statistical significance of within-group changes between round 1 and round 3 using the Wilcoxon signed rank test. We assessed the differences between the HC and IC groups using Mann-Whitney U tests. We used a Bonferroni adjusted alpha level of .025 in all statistical tests. We performed a thematic analysis to explore participants’ arguments during group discussion. Participants completed a satisfaction survey at the end of the consensus process.

Results: Of the 135 students completing a master of nursing and obstetrics, 120 participated in the experiment. We formed 8 HC groups (n=64) and 7 IC groups (n=56). The between-group comparison demonstrated that the human computation groups obtained a greater improvement in evidence scores compared to the IC groups, although the difference was not statistically significant. The between-group effect size was 0.56 (*P*=.30) for the medical imaging scenario, 0.07 (*P*=.97) for the therapeutic options scenario, and 0.89 (*P*=.11) for the drug use scenario. We found no significant differences in improvement in the degree of agreement between HC and IC groups. Between-group comparisons revealed that the HC groups showed greater improvement in degree of agreement for the medical imaging scenario (d=0.46, *P*=.37) and the drug use scenario (d=0.31, *P*=.59). Very few evidence arguments (6%) were quoted during informal group discussions.

Conclusions: Overall, the use of the IC method was appropriate as long as the evidence supported participants' beliefs or usual practice, or when the availability of the evidence was sparse. However, when some controversy about the evidence existed, the HC method outperformed the IC method. The findings of our study illustrate the importance of the choice of the consensus method in guideline development. Human computation could be an acceptable methodology for guideline development specifically for scenarios in which the evidence shows no resonance with participants' beliefs. Future research is needed to confirm the results of this study and to establish practical significance in a controlled setting of multidisciplinary guideline panels during real-life guideline development.

(J Med Internet Res 2013;15(1):e8) doi:10.2196/jmir.2055

KEYWORDS

clinical practice guidelines; evidence-based medicine; guideline development; consensus methods; human computation; games with a purpose

Introduction

Evidence-based clinical practice guidelines can be defined as "systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical conditions" [1]. They are intended to help physicians implement the burgeoning amount of scientific evidence on current medical best practices.

The development of clinical practice guidelines requires a systematic and transparent process, in which the recommendations of the clinical practice guideline (CPG) are explicitly linked to the clinical evidence. However, guidelines cannot be deduced from evidence alone, and expert opinions are needed to contextualize the evidence to the target population [2]. Furthermore, when evidence does not exist, or when there is little or incomplete evidence, the personal opinion of experts becomes more important [2]. Common methods for guideline development include the Delphi method, the nominal group technique (NGT), and the consensus development conference [3].

The high cost in time, resources, and efforts needed for the Delphi method, and the intensive commitment required for the NGT, pose important practical and logistic problems [3]. Various social-psychological influences on group discussion and decisions play an important role when face-to-face meetings or the NGT are used in the guideline development process. Previous research suggests that clinical evidence has a variable influence on guideline recommendations because of these social-psychological influences [4]. The variable influences of the clinical evidence, in turn, have an important impact on the validity and the quality of the guideline content as well as the implementation and effectiveness of the guideline [5,6].

Conducting the consensus development process entirely online is an approach that can address these concerns [7-10]. An online consensus process has the potential to involve a lot of participants and stakeholders, while offering organizational and logistic advantages in terms of cost and time savings. Social-psychological influences inherent in traditional face-to-face meetings could be eliminated by anonymously implementing the consensus process online. Explicit methods could be used to aggregate opinions.

Human-based computation is a technique in computer science in which the problems that a computer cannot yet solve are outsourced to humans. One class of human-based computation applications are the systems in which the tasks outsourced to humans are packaged as a game. These applications are called games with a purpose (GWAP). The idea behind these systems is to take advantage of people's desire to be entertained while performing useful tasks as a side effect. This approach is effective in building large knowledge databases, but to date it has no known practical applications in medicine [11-13].

Based on the principles of successful GWAP, we developed the CPGame (clinical practice game) application as a new method for guideline development. We built a prototype based on human-based computation techniques and the goals of the experiment. Our objectives were threefold: (1) to investigate the similarities or differences in degree of agreement and evidence with an informal consensus method to explore whether the human-based computation method is a valuable alternative, (2) to investigate arguments in decision making during group discussion, and (3) to explore perceptions and opinions about the consensus method. The objectives of the study were hypothesis-generating in the first place.

Methods

Design

We performed a randomized controlled trial (RCT) to compare the feasibility of the human-based computation method with an informal consensus method using a face-to-face meeting. Different consensus groups participated in the trial. Each of these groups consisted of 8 participants [14,15].

We developed 4 multiple choice scenarios involving lower back pain. An example scenario is included as Multimedia Appendix 1. Participants were asked to indicate their preference between several options for the given scenario. The 4 scenarios were totally different in content and concerned the following topics: (1) medical imaging, (2) therapeutic options, (3) drug use, and (4) sick leave. The first 3 scenarios included clear levels of clinical evidence for each of the answers participants could choose from. No clinical evidence was included for the fourth scenario. The evidence for the scenarios was selected based on a previous systematic review of the Belgian Health Care Knowledge Center [16]. The evidence was graded by the Belgian Health Care Knowledge Center using the GRADE system [17]. The quality of the evidence was classified as high, moderate, low, or very low.

Students completing a master of nursing and obstetrics at the University of Leuven participated in the experiment while taking a course in evidence-based health care. All of the participants already had a bachelor of nursing and practical experience. They all had baseline knowledge about lower back pain that was sufficient to judge the clinical scenarios. They also had sufficient knowledge in evidence-based health care to understand the evidence terms and the scientific meaning of the levels of evidence. There were no exclusion criteria, and all students had Internet experience.

Intervention and Controls

We developed the CPGame application based on human-based computation techniques. The human-based computation method is comparable to an online Delphi method packaged as a game. CPGame is a real-time collaborative application written in PHP, JavaScript, and Ajax. We used a MySQL database as the data repository. We pilot tested the technical robustness of the application with a group of trainees in family medicine.

Each participant in the HC groups was anonymously paired with another participant. The 2 participants played in a team against the other teams of 2 participants (giving 4 teams of 2 students in each group). Participants were given a user ID and password to log on to the CPGame application. The 2 participants on each team were given the same multiple choice questionnaire about a clinical scenario (Figure 1).

Independently of each other, both participants on a team gave their opinion about the case by choosing their preferred answer from the multiple choice list. When they had given their answers, the application displayed a message stating whether or not they reached consensus. If they reached consensus, they were given the next clinical scenario. If they did not reach a consensus, the application displayed the evidence for each answer and the answer of the teammate (Figure 2).

Each participant was given one chance to change his or her answer to try to reach consensus on a second attempt. After all the teams responded to all 4 scenarios, the CPGame application displayed the answers of all participants and the level of agreement between participants (Figure 3). Each participant reflected on his or her opinion and gave his or her final decision individually in a third round.

The game behind the human computation application consisted of a point system, a high-score list, and time pressure. These elements are described in the literature as being salient features that make GWAP fun [18]. The application determined team ranking based on the time in which a team completed all cases, the number of times consensus was reached, and the number of times the time limit was exceeded. Team ranking was only added as a competition element to make the consensus process more fun; it was not used as an outcome in final analysis. The third consensus round took place after team ranking was determined, to avoid a possible influence in the end results due to the competition element. The rules of the game were given to each participant on paper before the start of the game.

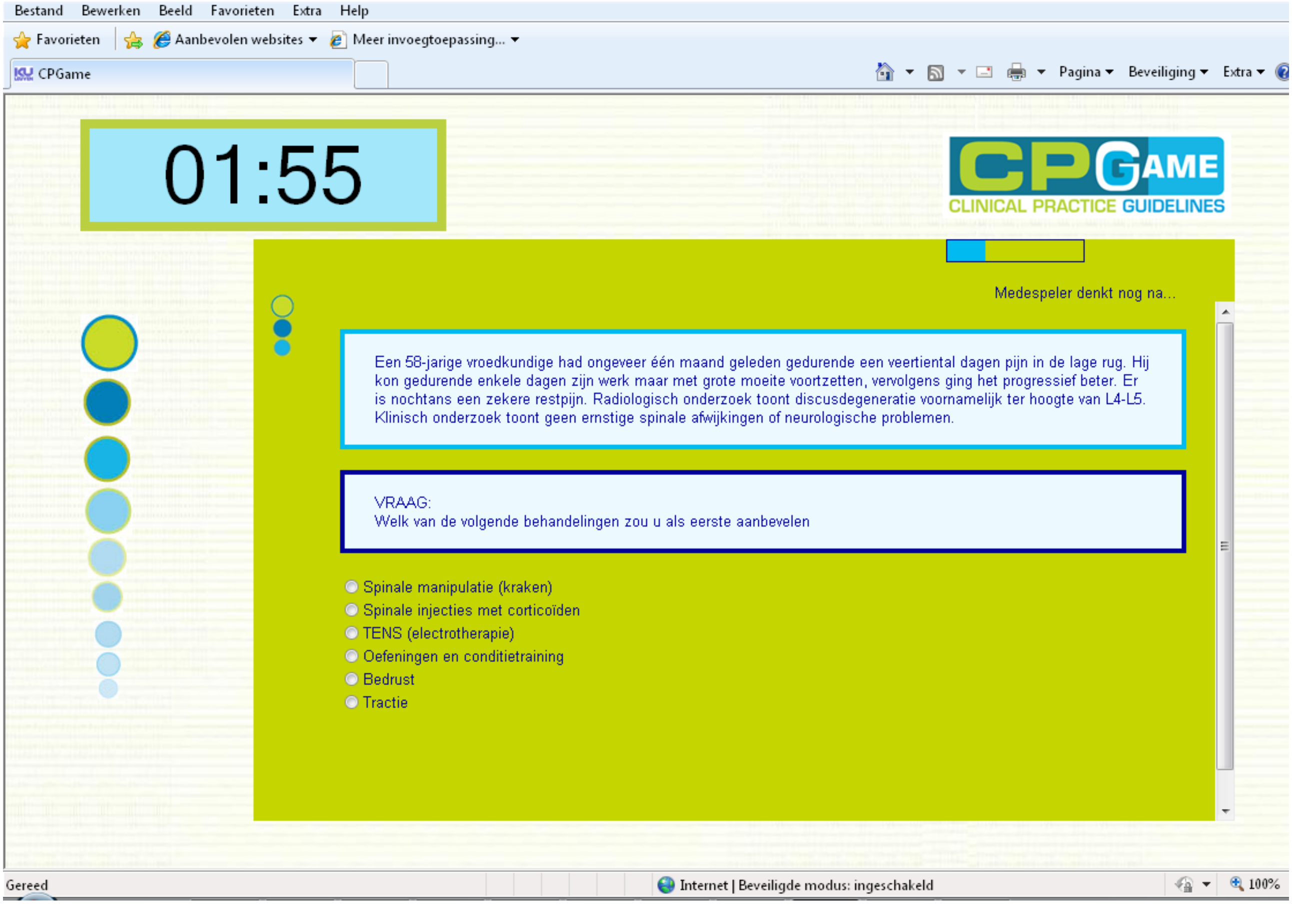
All participants completed the experiment in the same room at the same time. This one-room setting was possible because students participated in the trail during their course in evidence-based health care and had to come to the building for their class. A moderator was in the room in case there were technical problems during the experiment. The CPGame application was originally designed to be a self-directed process in which users independently participate online at their home.

The online approach to human-based computation (HC) the informal consensus method was comparable with a traditional face-to-face meeting. We took several measures to make both HC and IC methods comparable; the only difference was the mode (face-to-face versus Web-based packaged as a game).

The content of the scenarios and the evidence were similarly presented in both HC and IC groups. Before the start of the informal consensus process, each participant individually indicated his or her preference between several options for each scenario in the first round. After this first round, participants were randomly grouped into teams of 2, they were given the evidence for each treatment option, and each team discussed the cases. The discussion within the teams of 2 students was added as an additional step in the consensus process to ensure the points of measurement were equal between the 2 methods being reviewed. Each participant individually re-rated each scenario in the second round. After the second round, all 8 participants met in one group. A moderator, with experience in coordinating small groups, managed the discussion based on a written protocol designed to standardize the meetings. At the meeting, participants were told the judgments of the other group members and the research evidence. As a group, participants discussed each scenario and explored reasons for differences in opinions. After the group meeting, each participant individually re-rated each scenario in the third round.

Participants in neither the IC groups nor the HC groups were given incentives to participate in the experiment. The possible incentive of winning the game in the HC groups was neutralized by the introduction of a last consensus round outside the game.

Figure 1. Screen capture of a multiple choice scenario in the CPGame application.



01:55

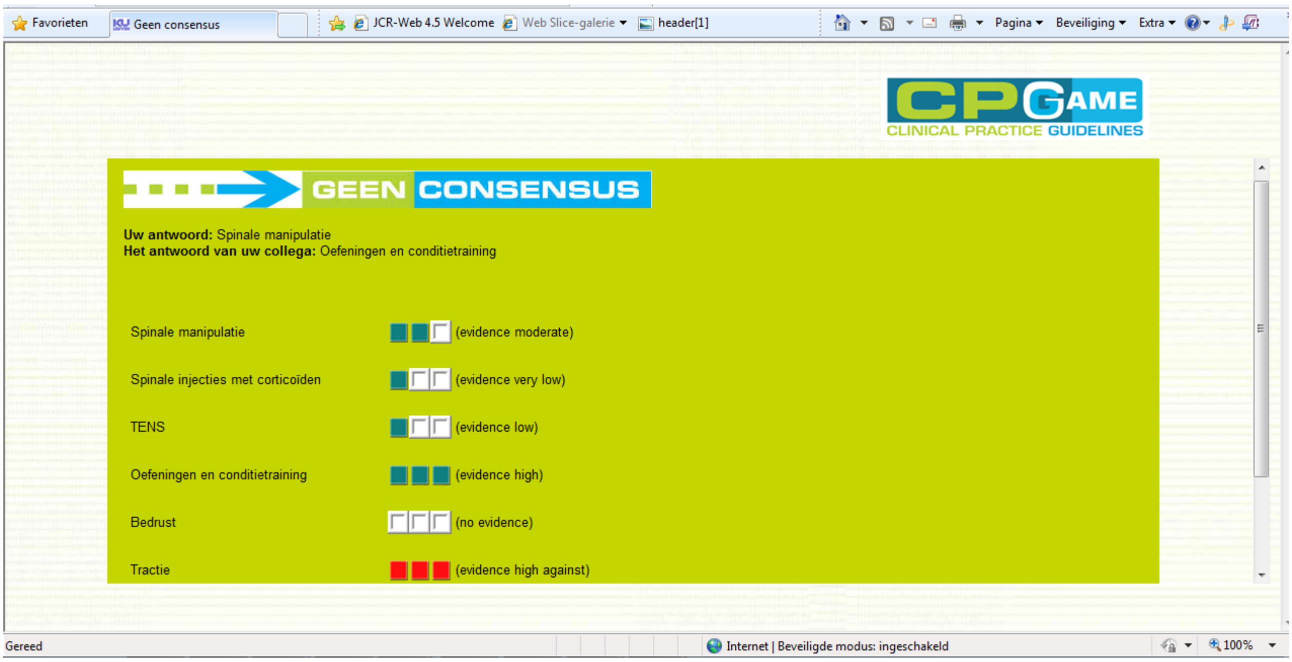
Medespeler denkt nog na...

Een 58-jarige vroedkundige had ongeveer één maand geleden gedurende een veertiental dagen pijn in de lage rug. Hij kon gedurende enkele dagen zijn werk maar met grote moeite voortzetten, vervolgens ging het progressief beter. Er is nochtans een zekere restpijn. Radiologisch onderzoek toont discusdegeneratie voornamelijk ter hoogte van L4-L5. Klinisch onderzoek toont geen ernstige spinale afwijkingen of neurologische problemen.

VRAAG:
Welk van de volgende behandelingen zou u als eerste aanbevelen

○ Spinale manipulatie (kraken)
○ Spinale injecties met corticoïden
○ TENS (electrotherapie)
○ Oefeningen en conditietraining
○ Bedrust
○ Tractie

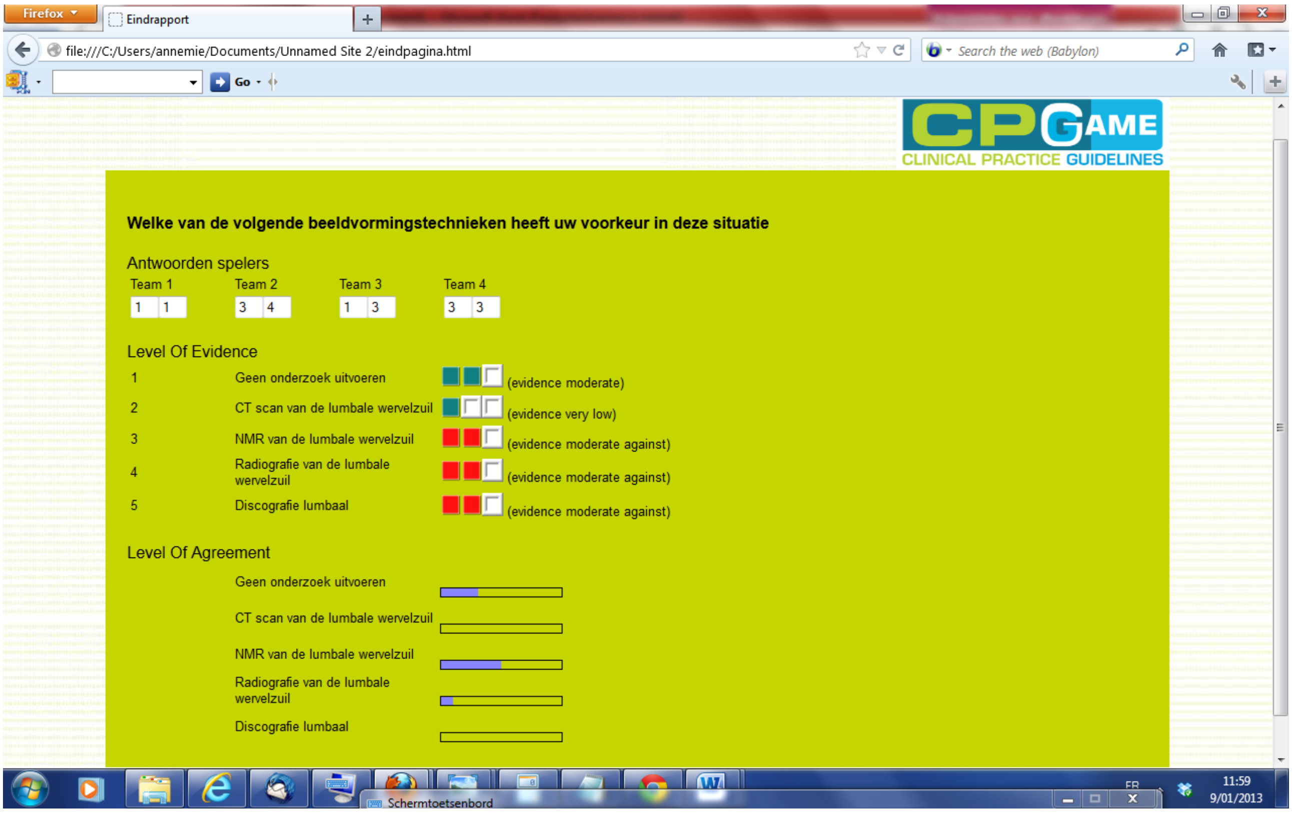
Figure 2. Screen capture of the information displayed on the CPGame when consensus was not reached within a team.



Uw antwoord: Spinale manipulatie
Het antwoord van uw collega: Oefeningen en conditietraining

Spinale manipulatie	■■□	(evidence moderate)
Spinale injecties met corticoïden	■□□	(evidence very low)
TENS	■□□	(evidence low)
Oefeningen en conditietraining	■■■	(evidence high)
Bedrust	□□□	(no evidence)
Tractie	🟥🟥🟥	(evidence high against)

Figure 3. Screen capture of the final page in the CPGame application, showing the answers of all participants, the level of evidence, and the level of agreement between participants.



Welke van de volgende beeldvormingstechnieken heeft uw voorkeur in deze situatie

Antwoorden spelers

Team 1	Team 2	Team 3	Team 4
1 1	3 4	1 3	3 3

Level Of Evidence

1	Geen onderzoek uitvoeren	🟦🟦☐	(evidence moderate)
2	CT scan van de lumbale wervelzuil	🟦☐☐	(evidence very low)
3	NMR van de lumbale wervelzuil	🟥🟥☐	(evidence moderate against)
4	Radiografie van de lumbale wervelzuil	🟥🟥☐	(evidence moderate against)
5	Discografie lumbaal	🟥🟥☐	(evidence moderate against)

Level Of Agreement

Geen onderzoek uitvoeren

CT scan van de lumbale wervelzuil

NMR van de lumbale wervelzuil

Radiografie van de lumbale wervelzuil

Discografie lumbaal

Randomization Procedure

All students were invited to participate in the trial a week before the experiment by one of the researchers. Participants were randomly assigned to 1 of the 2 groups (HC or IC) following simple randomization procedures. One of the researchers performed the randomization with an electronic random-list generator, initially in 1:1 ratio. A second step in the randomization procedure consisted of assigning individuals an additional consensus group number.

When they entered the computer room, participants of the human-based computation group blindly chose an envelope at random with a user ID and password to log on to the CPGame application. The envelope contained a number from 1 to 8. Numbers 1 and 2 played the game in a team, numbers 3 and 4 were a team, etc. Participants did not know each other's numbers and did not know who would be on each team. We used the same randomization procedure in the informal consensus group after first round ratings were completed. Participants blindly chose an envelope at random with a number from 1 to 8. Predetermined pairs of numbers were used to form the teams.

Group assignments were given just before the start of the experiment. Although participants knew they were participating in a guideline development project about lower back pain, they did not know the outcomes and the goals of the project before participation. Researchers were not blinded to allocation, but outcomes were objective measures.

Outcomes and Statistical Analyses

General

We conducted descriptive statistics and graphical displays to describe the sample population. Baseline data about the participants' gender and age were compared using chi-square and Mann-Whitney U tests, as appropriate.

There were 3 points of measurement: (1) before the consensus process (round 1), (2) during the consensus process (round 2), and (3) at the end of the consensus process (round 3). Primary outcomes focused on the change of opinions towards consensus and towards evidence between round 1 and round 3. An analysis of the group's level was warranted because of our interest in group decision making. As the group's outcomes were treated as individual observations, we had not taken the clustering of individuals within a discussion group into account. It was appropriate to analyze the 4 scenarios separately because they were totally different in content. As such, degree of (dis)agreement and degree of evidence were calculated for each of the clinical scenarios. We used a Bonferroni adjusted alpha level of .025 (.05/2 outcome measures for each clinical scenario) for all statistical tests. Predictive Analytics SoftWare Statistics 18 was used for statistical analyses.

Amount of (Dis)agreement

We used a kappa statistic to express the degree of (dis)agreement within a group at the different rounds. We estimated the within-group change between round 1 and round 3 by Cohen's d (calculated as the absolute difference between the kappa value at round 3 and the baseline kappa value at round 1, divided by the pooled standard deviation). We analyzed the statistical significance of within-group differences between round 1 and round 3 using Wilcoxon signed rank test.

We calculated between-group effect sizes for the differences in the change in agreement between the HC and IC groups to get an idea of the magnitude of the intervention effect on the amount of (dis)agreement. We calculated between-group effect sizes or Cohen's d as the absolute difference between the mean change in agreement across rounds in HC and the mean change in agreement across rounds in IC, divided by the pooled standard deviation. We assessed differences between the HC and IC groups in the change in degree of agreement using Mann-Whitney U tests. Final kappa scores were not adjusted for their baseline values because the subjects of comparison were the differences in the change in agreement across rounds, not the differences in final agreement.

Amount of Concordance with Clinical Evidence

We calculated a group's evidence score to have an idea of the degree of evidence in the answers of each group. We assigned different points to the different levels of evidence. An answer for which a high level of evidence existed got 4 points, a moderate level of evidence got 3 points, a low level of evidence got 2 points, and a very low level of evidence got 1 point. Answers for which there was evidence against got the same points with the opposite sign. Evidence points were multiplied by the number of participants who chose an answer with that level of evidence. The total sum was divided by the highest possible group's evidence score for the specific clinical question. An evidence score of 1 meant that all group members chose the answer with the highest level of evidence.

We estimated the within-group change in evidence score between round 1 and round 3 by Cohen's d (calculated as the absolute difference between the evidence score at round 3 and the baseline evidence score at round 1, divided by the pooled standard deviation). We analyzed the statistical significance of within-group differences between round 1 and round 3 using Wilcoxon signed rank test.

We calculated the between-group effect sizes for the differences in the change in evidence score between the HC and IC groups. We calculated the between-group effect sizes or Cohen's d as the absolute difference between the mean change in evidence scores across rounds in HC and the mean change in evidence scores across rounds in IC, divided by the pooled standard deviation. We assessed differences between the HC and IC groups in the change in evidence score using Mann-Whitney U tests.

Thematic Analysis

We used a hidden camera to record meetings of the informal consensus groups to explore arguments in each group's decision-making process. Hiding the camera was necessary to avoid social-psychological influences that arise with the awareness of recording. Two analysts anonymously transcribed and independently coded the recorded meetings of the face-to-face groups. Each communicative function within an utterance was defined as a dialogue act. Each dialogue act was coded and classified under a theme. We developed a preliminary list of themes based on the published list of themes created by Gardner et al [4]. We applied this preliminary list to the transcripts and adapted it to the specific situation of our populations.

We used the length of discussion time as a process measure of the group discussion. Discussion time was defined as the elapsed time between the start and the end of a group's decision-making activities.

We gave the students a paper questionnaire after the consensus process to explore perceptions about the consensus method. After the experiment, participants in the informal consensus group were notified about the hidden camera. Offline, we obtained informed consent to use the results for analysis from all participants. If one of the participants did not agree to allow us to use the hidden camera footage, we did not use the recordings and results of that group.

We obtained approval from the University Hospitals Leuven Medical Ethics Committee for this study in December 2009. The full protocol and the approval form can be obtained from the corresponding author.

Results

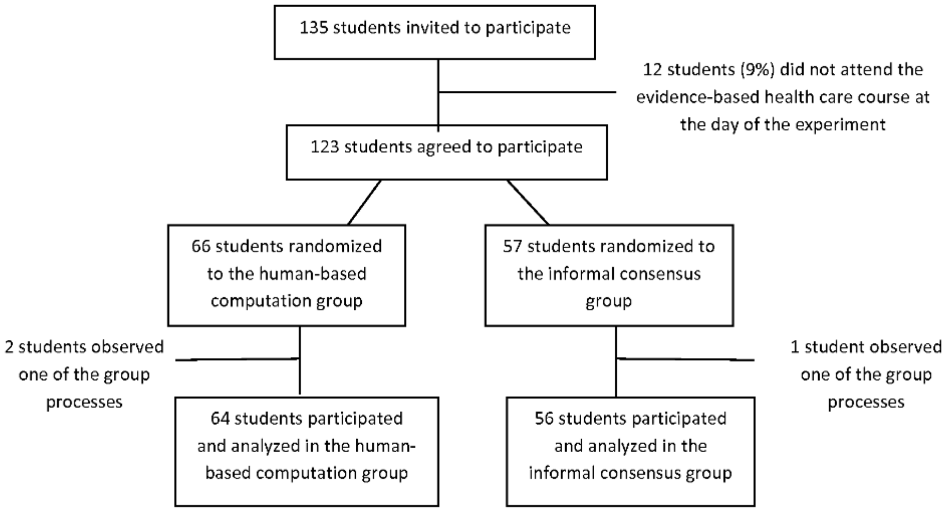
A total of 120 out of 135 students completing a master of nursing and obstetrics participated in the experiment. The participants formed 8 HC and 7 IC groups. Fewer students than expected attended the experiment, so only 7 instead of 8 informal consensus groups could be constituted. A total of 3 students were not assigned to groups. They participated as observers of the group's process and were not included in analysis. All participants who were randomly assigned to a group were analyzed in their original assigned groups (Figure 4).

The 2 groups were similar in terms of age and gender. There were no statistically significant differences in baseline evidence score and baseline agreement at round 1 (Table 1).

Table 1. Baseline demographic data and outcome scores.

	Human-based computation (8 groups, 64 participants)	Informal consensus (7 groups, 56 participants)
Year of birth	1986	1986
n (%) female	56 (88%)	50 (89%)
Evidence score		
Medical imaging (95% CI)	-0.15 (-0.40 to 0.10)	-0.14 (-0.32 to 0.04)
Therapeutic options (95% CI)	0.67 (0.51 to 0.82)	0.81 (0.72 to 0.89)
Drug use (95% CI)	0.43 (0.31 to 0.55)	0.49 (0.36 to 0.63)
Degree of agreement		
Medical imaging (95% CI)	0.29 (0.25 to 0.32)	0.29 (0.22 to 0.36)
Therapeutic options (95% CI)	0.45 (0.27 to 0.62)	0.5 (0.31 to 0.69)
Drug use (95% CI)	0.21 (0.11 to 0.31)	0.25 (0.11 to 0.38)
Sick leave (95% CI)	0.52 (0.37 to 0.66)	0.60 (0.34 to 0.85)

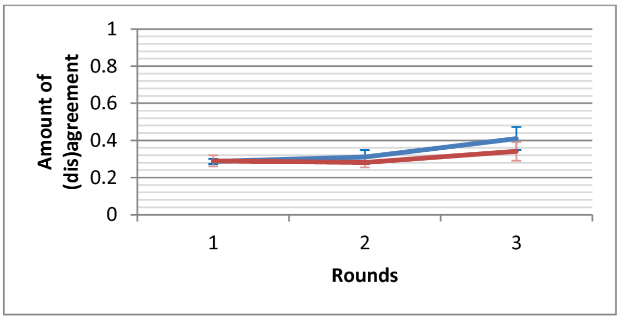
Figure 4. Flowchart showing participants in the trial.



Amount of (Dis)agreement

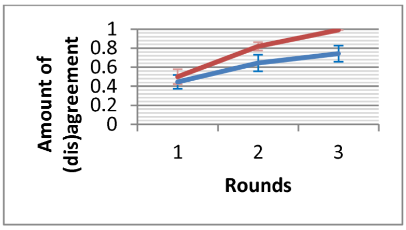
Within-group effect sizes (Cohen’s d) varied between 0.26 and 2.53 in the HC groups and were statistically significant for the therapeutic options scenario (d=1.44 with *P*=.02) and for the drug use scenario (d=2.53 with *P*=.01). Within-group effect sizes in the IC groups varied between 0.39 and 2.33. IC groups showed a significant improvement in the degree of agreement for the therapeutic options scenario (d=2.33 with *P*=.02). No significant differences in improvement of degree of agreement were found between HC and IC groups. Between-group comparisons revealed that the HC groups showed greater improvement in degree of agreement for the medical imaging scenario (d=0.46 with *P*=.37) and the drug use scenario (d=0.31 with *P*=.59). The opposite was true for the therapeutic options scenario (d=-0.9 with *P*=.10) and the sick leave scenario (d=-1.25 with *P*=.05). The change in degree of agreement across the 3 rounds, within-group effect sizes, and between-group effect sizes are displayed in Figures 5 to 8.

Figure 5. Amount of (dis)agreement for the medical imaging scenario.



	Round 1 (95% CI)	Round 2 (95%CI)	Round 3 (95%CI)	Within-group effect size	*P*	Between-group effect size	*P*
HC (n = 8)	0.29 (0.25 to 0.32)	0.31 (0.22 to 0.40)	0.41 (0.26 to 0.56)	0.69	.04		
						0.46	.37
IC (n = 7)	0.29 (0.22 to 0.36)	0.28 (0.22 to 0.35)	0.34 (0.22 to 0.46)	0.39	.27		

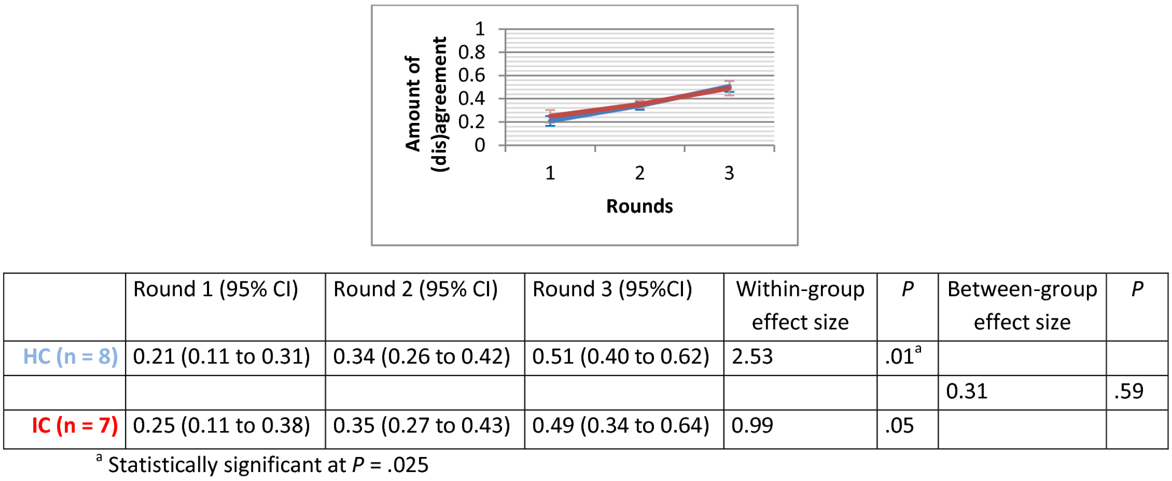
Figure 6. Amount of (dis)agreement for the therapeutic options scenario.



	Round 1 (95% CI)	Round 2 (95% CI)	Round 3 (95% CI)	Within-group effect size	*P*	Between-group effect size	*P*
HC (n = 8)	0.45 (0.27 to 0.62)	0.65 (0.44 to 0.85)	0.74 (0.54 to 0.94)	1.44	.02[a]		
						-0.90	.10
IC (n = 7)	0.5 (0.31 to 0.69)	0.82 (0.71 to 0.93)	0.99 (0.99 to 0.99)	2.33	.02[a]		

[a] Statistically significant at *P* = .025

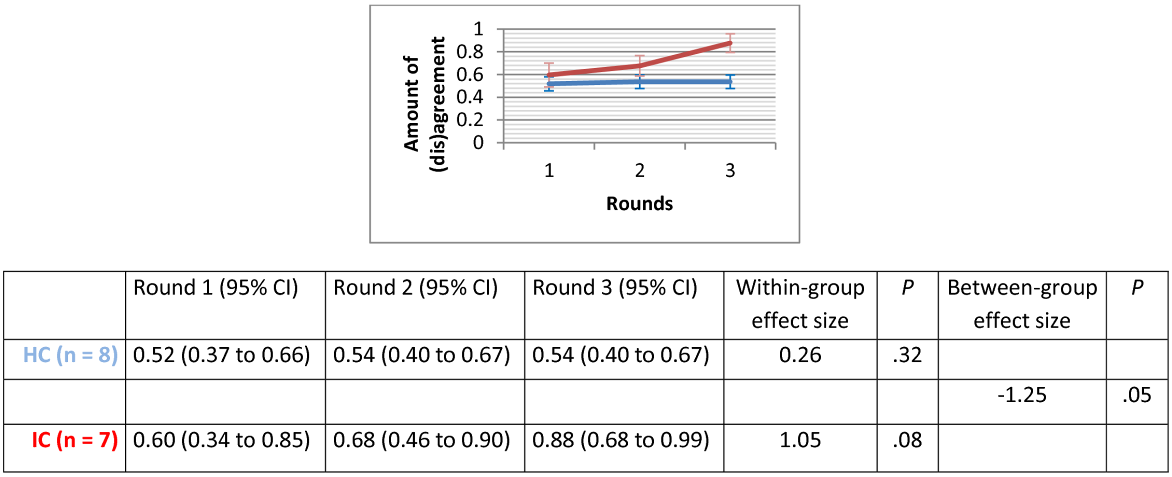
Figure 7. Amount of (dis)agreement for the drug use scenario.



	Round 1 (95% CI)	Round 2 (95% CI)	Round 3 (95%CI)	Within-group effect size	*P*	Between-group effect size	*P*
HC (n = 8)	0.21 (0.11 to 0.31)	0.34 (0.26 to 0.42)	0.51 (0.40 to 0.62)	2.53	.01[a]		
						0.31	.59
IC (n = 7)	0.25 (0.11 to 0.38)	0.35 (0.27 to 0.43)	0.49 (0.34 to 0.64)	0.99	.05		

[a] Statistically significant at *P* = .025

Figure 8. Amount of (dis)agreement for the sick leave scenario.

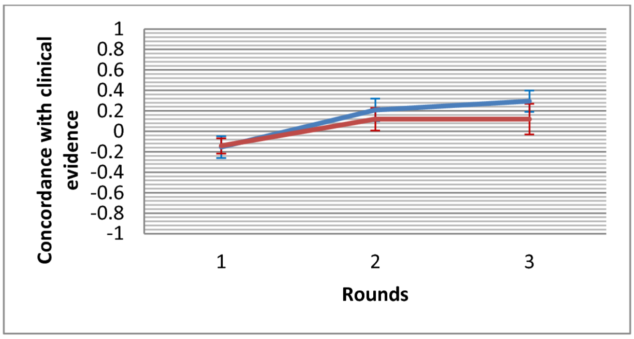


	Round 1 (95% CI)	Round 2 (95% CI)	Round 3 (95% CI)	Within-group effect size	*P*	Between-group effect size	*P*
HC (n = 8)	0.52 (0.37 to 0.66)	0.54 (0.40 to 0.67)	0.54 (0.40 to 0.67)	0.26	.32		
						-1.25	.05
IC (n = 7)	0.60 (0.34 to 0.85)	0.68 (0.46 to 0.90)	0.88 (0.68 to 0.99)	1.05	.08		

Concordance with Clinical Evidence

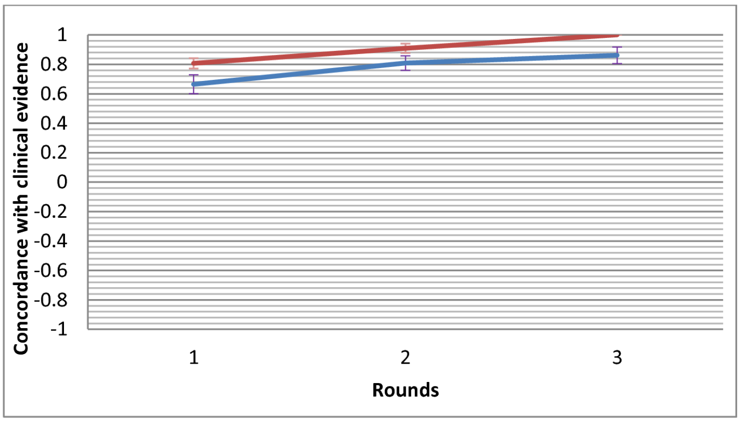
After 3 rounds, the mean evidence score increased for all clinical scenarios in both groups. Within-group changes showed a significant improvement in evidence score for the drug use scenario in the HC groups (d=3.67 with *P*=.01) and for the therapeutic options scenario in the IC groups (d=2.11 with *P*=.02). The between-group comparison demonstrated that the human-based computation groups obtained a greater improvement in evidence scores compared to the IC groups, although the difference was not statistically significant. Between-group effect size was 0.56 (*P*=.30) for the medical imaging scenario, 0.07 (*P*=.97) for the therapeutic options scenario and 0.89 (*P*=.11) for the drug use scenario. Figures 9 to 11 show the change in mean group's evidence score across the 3 rounds, within-group effect sizes and between-group effect sizes for the different clinical scenarios.

Figure 9. Concordance with clinical evidence for the medical imaging scenario.



	Round 1 (95% CI)	Round 2 (95% CI)	Round 3 (95% CI)	Within-group effect size	*P*	Between-group effect size	*P*
HC (n = 8)	-0.15 (-0.40 to 0.10)	0.21 (-0.05 to 0.47)	0.29 (0.05 to 0.54)	1.07	.04		
						0.56	.30
IC (n = 7)	-0.14 (-0.32 to 0.04)	0.12 (-0.15 to 0.39)	0.12 (-0.25 to 0.49)	1.06	.05		

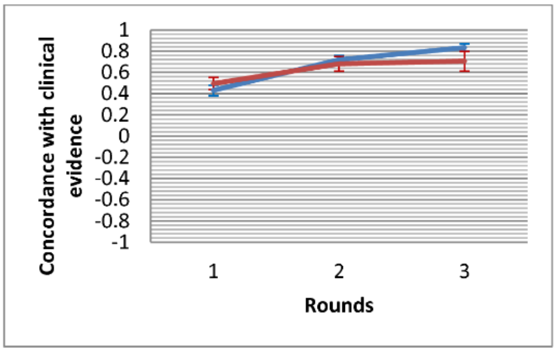
Figure 10. Concordance with clinical evidence for the therapeutic options scenario.



	Round 1 (95% CI)	Round 2 (95% CI)	Round 3 (95% CI)	Within-group effect size	*P*	Between-group effect size	*P*
HC (n = 8)	0.67 (0.51 to 0.82)	0.81 (0.69 to 0.93)	0.86 (0.73 to 0.99)	1.03	.03		
						0.07	.97
IC (n = 7)	0.81 (0.72 to 0.89)	0.91 (0.83 to 0.99)	1.00 ()	2.11	.02[b]		

[b] Statistically significant at *P* = .025

Figure 11. Concordance with clinical evidence for the drug use scenario.



	Round 1 (95% CI)	Round 2 (95% CI)	Round 3 (95% CI)	Within-group effect size	*P*	Between-group effect size	*P*
HC (n = 8)	0.43 (0.31 to 0.55)	0.72 (0.62 to 0.81)	0.83 (0.75 to 0.91)	3.67	.01[b]		
						0.89	.11
IC (n = 7)	0.49 (0.36 to 0.63)	0.68 (0.51 to 0.85)	0.70 (0.48 to 0.93)	0.67	.09		

[b] Statistically significant at *P* = .025

Thematic Analysis

Figure 12 shows the frequencies with which the different themes appeared across the meetings. The results show a greater focus on clinical preference than on clinical evidence. Themes relating to clinical judgment or preference occurred the most (177/369, 48% arguments), while there were relatively few arguments explicitly pro evidence (23/369, 6% arguments). Group discussions were characterized by a high degree of uncertainty, which is shown by the high percentages (87/369, 24%) in category V, and few references to own or other's clinical experiences (23/369, 6%). Only 5 out of 369 agreements (1%) could be classified under the category “reference to other guidelines or literature.” All individual meetings followed more or less the same patterns in themes. A mean Cohen’s kappa of 0.77 was reached for interanalyst agreement.

Mean discussion time for the 4 clinical scenarios was 32.9 minutes (± 6.5 minutes) in the IC groups and 14.6 minutes (± 2.2 minutes) in the HC groups. Analysis of the time intervals revealed a statistically significant shorter discussion time in the HC groups compared to the IC groups (*P*=.001). Participant satisfaction scores are shown in Table 2 and 3.

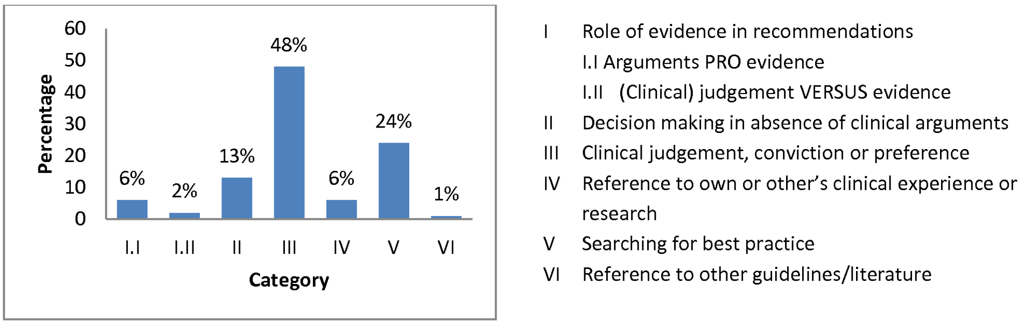
Table 2. Satisfaction scores.

		Human-based computation method (HC) n (%)	Informal consensus method (IC) n (%)
I am satisfied with the group answer	Strongly agree	7 (11)	12 (21)
	Agree	30 (48)	30 (53)
	Undecided	21 (33)	11 (20)
	Disagree	5 (8)	2 (4)
	Strongly disagree	0 (0)	1 (2)
		63 (100)	56 (100)
I find the levels of evidence important when making my decision	Strongly agree	32 (50)	18 (32)
	Agree	25 (39)	32 (57)
	Undecided	6 (9)	5 (9)
	Disagree	1 (1)	1 (2)
	Strongly disagree	0 (0)	0 (0)
		64 (100)	56 (100)

Table 3. Decision-making scores.

I would describe the decision-making process as:						
Efficient	1	2	3	4	5	Not efficient
	n (%)	n (%)	n (%)	n (%)	n (%)	
HC	6 (10)	30 (48)	14 (23)	12 (19)	0 (0)	
IC	11 (20)	27 (48)	15 (27)	1 (2)	2 (4)	

Figure 12. Percentage of arguments in the different categories of the coding scheme.



Discussion

Principal Results

For the cases with evidence, changes in answers across rounds were more evidence-based in the HC groups compared to the IC groups. HC groups obtained a greater improvement in evidence scores compared to the IC groups. The anonymity of the participants in the HC game evidently avoided direct social-psychological influencing, as intended.

Differences in the improvement in agreement across rounds were better in the HC groups for the medical imaging scenario and the drug use scenario, but not for the therapeutic options scenario. The evidence score for that scenario was already relatively high starting at round 1 in the IC groups. The evidence supported students' beliefs, values, and preexisting opinions and little group pressure was needed to convince a few individuals to reach full consensus in the IC groups.

For the sick leave scenario, which did not include clinical evidence, the informal consensus (IC) groups demonstrated closer group agreement compared to the human computation (HC) groups. Opinions were more likely to shift when groups met face-to-face, as suggested by the study of Hutchings et al [19]. The choice of the degree of (dis)agreement as a process measure assumed that consensus is a good outcome and that IC groups fared better for the scenario without evidence. Many guideline developers would disagree with the fact that consensus is a good outcome. Although we acknowledge this point of view, we believe it was appropriate to use kappa values as a process measure because reaching consensus is the primary goal of each consensus process.

Supplying the evidence at round 2 had an influence on group judgment (shown by the positive within-group Cohen's d for the evidence score) as well in the HC groups as in the IC groups. However, thematic analysis in the IC groups revealed that choices were more likely to be based on clinical judgment or conviction, rather than on clinical evidence (as supported by Raine [20]). The few evidence arguments in the IC groups (6%) were in sharp contrast with the results of the questionnaire, where 89% of the participants in the IC groups perceived the influence of the evidence as important. Perceptions did not correspond with arguments used in practice.

Hutching et al [19] demonstrated earlier that direct exposure to arguments and (dominating) personalities could lead guideline development groups in different directions. The anonymity of the participants in the HC groups eliminated important aspects of social-psychological influences, which gained the upper hand in the IC groups. It is surprising that so little evidence arguments were quoted in the IC groups. These findings could confirm the idea of Sauerland et al [21] that evidence in IC groups is used to confirm preexisting opinions, rather than to change them.

Limitations of the Study

There were some methodological and practical limitations to our study. The limited number of clinical scenarios, especially for the type of case without evidence, may reduce the generalization of the results. It was a proof-of-concept hypothesis-generating study, so we did not power the study before the start. The obtained power was not robust enough for a reliable detection of a between-group effect, which increased our chance of false-negative conclusions concerning statistical significance.

The high degree of variability between the individual groups may be seen as realistic reflections of variations in clinical perspective. However, it also confirms the importance of the composition of the guideline panel and the choice of the moderator [22-25]. What is to be decided is often already determined with the selection of the deciders [26]. The effect of the panel composition should be minimal because it concerned a rather homogeneous population (students in an evidence-based heath care course) randomized to the different groups (HC or IC). Although in contrast with real multidisciplinary guideline panels, we chose a rather homogeneous group of students with limited expertise as the subject of the experiment to partly control for the social-psychological influences rising from multistatus groups. The choice of the study population was appropriate based on the research, which explored the influence of the consensus method on the change of opinions towards consensus and evidence, rather than the content of consensus. We decided that a high degree of lower back pain specialization was not necessary in this preliminary phase of the research. Because of the exploratory and early phase nature of the work, the choice of these participants with baseline knowledge about lower back pain was justified. We stressed spontaneous group interaction, rather than reaching consensus, to minimize the influence of the moderator.

Time intervals did not represent real discussion times in multidisciplinary guideline panels because of the aforementioned differences between our discussion groups and these expert panels. However, time analysis demonstrated the potential of the human-based computation method to be an efficient consensus method. The thematic analyses may also differ from discussions in expert groups or other guideline development groups. Participants were not specialized in one medical domain and did not have the intention to over-state the effectiveness of their specialist intervention, which may have reduced the amount of contentious issues. The thematic analyses also reflected the attitude towards clinical practice of a new generation of professionals recently educated in evidence-based health care.

We gave equal weighting to the different levels of evidence in our calculation of the evidence score. We are aware that not all people give equal weights to a difference between evidence low and evidence moderate, or evidence moderate and evidence high, etc. However, the literature did not provide a scientific basis to assign unequal weights to the different evidence levels. As a consequence, equal weighting of the different evidence levels was considered as appropriate as unequal weighting.

Another limitation of the study is the lack of a third online Delphi group without a game component to allow us to separately study the effect of the online approach and the effect of the game itself. While eliminating social-psychological influences inherent in face-to-face groups, we introduced a new psychological element of competition in the HC groups due to the game component. This could have influenced the results at round 2, but did not affect our final results. Participants got the chance to reconsider their answers in a third individual round after finishing the game.

The students participated in the experiment during the hours of their class in evidence-based health care. Because of this, we could easily conduct the experiment with all the participants in one room. No additional logistic facilities were required; students had to come to the building for their class. This is in contrast to the suggested advantages of an online process, where experts participate at home. However, the method was originally designed to be a self-directed process, participants worked individually on a computer during the experiment, and no verbal communication with the moderator took place. Therefore, we believe the results could be easily generalized to a real-world online process.

Relation to Other Studies

Relatively few studies compared different consensus methods for guideline development. Many of them differed in the consensus methods they compared or in the way they operationalized the method. Washington et al [9] and Kadam et al [22] did not demonstrate differences in final ratings between the consensus methods. Shekelle et al found limited differences in their study [27]. Hutching et al [19] showed greater within-group agreement in nominal groups compared with Delphi groups, which was contrary to the earlier research of Leape et al [28]. The systematic review of Murphy et al [3] concludes that formal methods generally perform better than informal ones and may be better for consensus development.

Our newly developed method of guideline development by human-based computation proved very useful in the introduction of clinical evidence arguments, while neutralizing for social-psychological influences by authoritarian opinions.

The findings of our study illustrate once more the importance of the choice of the consensus method in guideline development. Giving the same evidence summary and using a consensus process, HC and IC groups could come to different group views. The influence of the consensus method seemed to depend on the type of clinical question. Overall, the use of the informal consensus method may be appropriate as long as the evidence supports participants’ beliefs or usual practice, or when the availability of the evidence is sparse. However, when some controversy about the evidence exists, one could doubt the appropriateness of the informal consensus method. Because guideline programs are intended to reduce inappropriate variations in health care, guidelines are more important for clinical questions where the evidence shows no resonance with participants’ beliefs. Human computation outperformed the informal consensus method for this type of clinical questions.

Human Computation: an Acceptable Method for Guideline Development?

Participants perceived the human-based computation method as pleasant and enjoyable. Satisfaction was similar in the HC and IC groups. Only 3 out of 56 participants (5%) were dissatisfied with the group answer in the IC groups, while 5 out of 63 participants (8%) were dissatisfied in the HC groups. This is in contrast to the literature on computer-mediated communication versus face-to-face groups [10,29], where lower satisfaction is reported in computer groups in general. The additional game component in the computer groups could be a possible explanation for these higher satisfaction levels.

However, the major strengths of the method (the anonymity of panelists, the elimination of social-psychological influences in face-to-face meetings, and the possibility to participate in an online development group from a distance) was at the same time a reason for lower satisfaction. Participants perceived lack of group discussion and interaction in the HC groups as a negative aspect of the method. Participants seemed to need the opportunity to find out reasons for other members' decisions [3]. Although the human-based computation method has the potential to offer advantages in terms of logistics and more objective decision making, participants perceived the method's efficiency more negatively in the human-based computation group than in the informal consensus group. Of the 62 in the HC groups participants, 12 (19%) were dissatisfied with the efficiency of the HC method, while 3 out of 56 participants in the IC groups (5%) were dissatisfied with the efficiency of the informal consensus method. This was probably related to the lack of understanding of other participants' arguments.

The current format of the CPGame application was built for the purposes of the experiment. This prototype was essential to test the feasibility and the acceptability of the specific method for guideline development. However, if it is to be useful in practice, a more complex application will be needed.

We believe it is important to draw on the advantages of both methods (human-based computation and face-to-face meetings) in view of future system improvements. The exploration of group views should be incorporated, while maintaining the existing advantages of human-based computation. A hybrid method could be considered, including an extra button to ask for the arguments of other players to complement the human computation method. An extension to an asynchronous mode would also allow large-scale advantages and let people choose when they participate in the process. We chose the current format of multiple choice questions to test the feasibility of the method because of its plainness. Extensions to other question formats are also feasible and probably more adapted to guideline development.

Conclusions

The findings of our study illustrate the importance of the choice of the consensus method in guideline development. Giving the same evidence summary and using a different consensus process, two groups can come to different group views, which implies a considerable risk towards conflicting guideline recommendations on the same topic.

Human computation could be a time efficient and acceptable methodology for guideline development specifically for scenarios in which the evidence shows no resonance with participants' beliefs. Changes in evidence scores and agreement after 3 rounds were higher in HC groups compared to IC groups for this type of scenario. Controlled feedback is given while eliminating the social-psychological components of a group process. Level of evidence and level of agreement are separated, which could increase transparency of the guideline-development process.

Future research is needed to confirm the results and to establish practical significance in a controlled setting of multidisciplinary guideline panels during real-life guideline development.

Acknowledgments

We are grateful to Professor Dr G. Verbeke, Professor Dr M. Euwema, and Professor B. De Moor and his research group for their valuable advice. We are thankful to Jody Pauwels for his practical assistance. We would like to thank all the participants who took part in the experiment.

Conflicts of Interest

The authors and evaluators are the developers of CPGame, a human-based computation application. There were no other conflicts of interest.

Multimedia Appendix 1

Example of a multiple-choice questionnaire based on a clinical scenario involving lower back pain.

[PDF File (Adobe PDF File), 9KB - jmir_v15i1e8_app1.pdf]

Multimedia Appendix 2

CONSORT E-health checklist V1.6.1 [30].

[PDF File (Adobe PDF File), 572KB - jmir_v15i1e8_app2.pdf]

References

1. Commitee to Advise the Public Health Service on Clinical Practice Guidelines, Institute of Medicine. In: Field MJ, Lohr KN, editors. Clinical Practice Guidelines: Directions for a New Program. Washington, DC: National Academies Press; 1990.

2. Shekelle PG, Woolf SH, Eccles M, Grimshaw J. Clinical guidelines: developing guidelines. BMJ 1999 Feb 27;318(7183):593-596 [FREE Full text] [Medline: 10037645]
3. Murphy MK, Black NA, Lamping DL, McKee CM, Sanderson CF, Askham J, et al. Consensus development methods and their use in clinical guideline development. Health Technol Assess 1998;2(3):i-iv, 1 [FREE Full text] [Medline: 9561895]
4. Gardner B, Davidson R, McAteer J, Michie S. A method for studying decision-making by guideline development groups. Implement Sci 2009;4:48 [FREE Full text] [doi: 10.1186/1748-5908-4-48] [Medline: 19656366]
5. Pagliari C, Grimshaw J, Eccles M. The potential influence of small group processes on guideline development. J Eval Clin Pract 2001 May;7(2):165-173. [Medline: 11489041]
6. Sheldon TA, Cullum N, Dawson D, Lankshear A, Lowson K, Watt I, et al. What's the evidence that NICE guidance has been implemented? Results from a national evaluation using time series analysis, audit of patients' notes, and interviews. BMJ 2004 Oct 30;329(7473):999 [FREE Full text] [doi: 10.1136/bmj.329.7473.999] [Medline: 15514342]
7. Bowles KH, Holmes JH, Naylor MD, Liberatore M, Nydick R. Expert consensus for discharge referral decisions using online delphi. AMIA Annu Symp Proc 2003:106-109 [FREE Full text] [Medline: 14728143]
8. Deshpande AM, Shiffman RN. Delphi rating on the Internet. AMIA Annu Symp Proc 2003:828 [FREE Full text] [Medline: 14728333]
9. Washington DL, Bernstein SJ, Kahan JP, Leape LL, Kamberg CJ, Shekelle PG. Reliability of clinical guideline development using mail-only versus in-person expert panels. Med Care 2003 Dec;41(12):1374-1381. [doi: 10.1097/01.MLR.0000100583.76137.3E] [Medline: 14668670]
10. Fjermestad J, Hiltz SR. Group support systems: a descriptive evaluation of case and field studies. JMIS 2001;17(3):115-159.
11. Lieberman H, Smith D, Teeters A. Common consensus: a web-based game for collecting commonsense goals. 2007 Jan Presented at: Proceedings of Intelligent User Interfaces (IUI'07); 2007; Honolulu, Hawaii.
12. Von Ahn L, Dabbish L. Labeling images with a computer game. 2004 Apr Presented at: Proceedings of the SIGCHI conference on human factors in computing systems (CHI'04); 2004; Vienna, Austria p. 24-29. [doi: 10.1145/985692.985733]
13. Von Ahn L, Kedia M, Blum M. Verbosity: a game for collecting common sense facts. 2006 Apr Presented at: Proceedings of the SIGCHI conference on human factors in computing systems (CHI'06); 2006; Montreal, Canada. [doi: 10.1145/1124772.1124784]
14. Richardson FM. Peer review of medical care. Med Care 1972;10(1):29-39. [Medline: 5007750]
15. Black N, Murphy M, Lamping D, McKee M, Sanderson C, Askham J, et al. Consensus development methods: a review of best practice in creating clinical guidelines. J Health Serv Res Policy 1999 Oct;4(4):236-248. [Medline: 10623041]
16. Nielens H, Van Zundert J, Mairiaux P, Gailly J, Van Den Hecke N, Mazina D. Chronische lage rugpijn. In: Good Clinical Practice (GCP). Brussel: Federaal Kenniscentrum voor de gezondheidszorg (KCE); 2006.
17. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008 Apr 26;336(7650):924-926 [FREE Full text] [doi: 10.1136/bmj.39489.470347.AD] [Medline: 18436948]
18. Von Ahn L. Games with a purpose. Computer 2006;39(6):92-94. [doi: 10.1109/MC.2006.196]
19. Hutchings A, Raine R, Sanderson C, Black N. A comparison of formal consensus methods used for developing clinical guidelines. J Health Serv Res Policy 2006 Oct;11(4):218-224. [doi: 10.1258/135581906778476553] [Medline: 17018195]
20. Raine R, Sanderson C, Hutchings A, Carter S, Larkin K, Black N. An experimental study of determinants of group judgments in clinical guideline development. Lancet 2004;364(9432):429-437. [doi: 10.1016/S0140-6736(04)16766-4] [Medline: 15288741]
21. Sauerland S, Neugebauer E. Consensus conferences must include a systematic search and categorization of the evidence. Surg Endosc 2000 Oct;14(10):908-910. [Medline: 11080401]
22. Kadam UT, Jordan K, Croft PR. A comparison of two consensus methods for classifying morbidities in a single professional group showed the same outcomes. J Clin Epidemiol 2006 Nov;59(11):1169-1173. [doi: 10.1016/j.jclinepi.2006.02.016] [Medline: 17027427]
23. Fretheim A, Schünemann HJ, Oxman AD. Improving the use of research evidence in guideline development: 3. Group composition and consultation process. Health Res Policy Syst 2006;4:15 [FREE Full text] [doi: 10.1186/1478-4505-4-15] [Medline: 17134482]
24. Carpenter J, Hutchings A, Raine R, Sanderson C. An experimental study of the influence of individual participant characteristics on formal consensus development. Int J Technol Assess Health Care 2007;23(1):108-115. [doi: 10.1017/S0266462307051641] [Medline: 17234024]
25. Campbell SM, Hann M, Roland MO, Quayle JA, Shekelle PG. The effect of panel membership and feedback on ratings in a two-round Delphi survey: results of a randomized controlled trial. Med Care 1999 Sep;37(9):964-968. [Medline: 10493474]
26. Sniderman AD, Furberg CD. Why guideline-making requires reform. JAMA 2009 Jan 28;301(4):429-431. [doi: 10.1001/jama.2009.15] [Medline: 19176446]
27. Shekelle PG, Schriger DL. Evaluating the use of the appropriateness method in the Agency for Health Care Policy and Research Clinical Practice Guideline Development process. Health Serv Res 1996 Oct;31(4):453-468 [FREE Full text] [Medline: 8885858]

28. Leape LL, Freshour MA, Yntema D, Hsiao W. Small-group judgment methods for determining resource-based relative values. Med Care 1992 Nov;30(11 Suppl):NS28-NS39. [Medline: 1434965]
29. Sarrina Li S. Computer-mediated communication and group decision making: a functional perspective. Small Group Research 2007;28:593. [doi: 10.1177/1046496407304335]
30. 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

CPG: clinical practice guideline
CPGame: clinical practice game
GWAP: games with a purpose
HC: human-based computation
IC: informal face-to-face consensus method
NGT: nominal group technique
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 18.01.12; peer-reviewed by S Gupta, A Klovning; comments to author 25.06.12; revised version received 13.07.12; accepted 23.09.12; published 17.01.13.

Please cite as:
Heselmans A, Aertgeerts B, Donceel P, Van de Velde S, Vanbrabant P, Ramaekers D
Human Computation as a New Method for Evidence-Based Knowledge Transfer in Web-Based Guideline Development Groups: Proof of Concept Randomized Controlled Trial
J Med Internet Res 2013;15(1):e8
URL: http://www.jmir.org/2013/1/e8/
doi:10.2196/jmir.2055
PMID:23328663

Original Paper

Cardiopulmonary Resuscitation Training in High School Using Avatars in Virtual Worlds: An International Feasibility Study

Johan Creutzfeldt[1], MD; Leif Hedman[1,2], PhD; LeRoy Heinrichs[3], MD, PhD; Patricia Youngblood[4], PhD; Li Felländer-Tsai[1], MD, PhD

[1]Department of Clinical Science, Intervention and Technology, Karolinska Institutet, Stockholm, Sweden
[2]Department of Psychology, Umeå University, Umeå, Sweden
[3]Innovation in Learning, Inc., Los Altos Hills, CA, United States
[4]Division of Clinical Anatomy, Department of Surgery, Stanford University School of Medicine, Palo Alto, CA, United States

Corresponding Author:
Johan Creutzfeldt, MD
Department of Clinical Science, Intervention and Technology
Karolinska Institutet
K32
Stockholm, 14186
Sweden
Phone: 46 858580000 ext 86334
Fax: 46 858582224
Email: johan.creutzfeldt@ki.se

Abstract

Background: Approximately 300,000 people suffer sudden cardiac arrest (SCA) annually in the United States. Less than 30% of out-of-hospital victims receive cardiopulmonary resuscitation (CPR) despite the American Heart Association training over 12 million laypersons annually to conduct CPR. New engaging learning methods are needed for CPR education, especially in schools. Massively multiplayer virtual worlds (MMVW) offer platforms for serious games that are promising learning methods that take advantage of the computer capabilities of today's youth (ie, the digital native generation).

Objective: Our main aim was to assess the feasibility of cardiopulmonary resuscitation training in high school students by using avatars in MMVM. We also analyzed experiences, self-efficacy, and concentration in response to training.

Methods: In this prospective international collaborative study, an e-learning method was used with high school students in Sweden and the United States. A software game platform was modified for use as a serious game to train in emergency medical situations. Using MMVW technology, participants in teams of 3 were engaged in virtual-world scenarios to learn how to treat victims suffering cardiac arrest. Short debriefings were carried out after each scenario. A total of 36 high school students (Sweden, n=12; United States, n=24) participated. Their self-efficacy and concentration (task motivation) were assessed. An exit questionnaire was used to solicit experiences and attitudes toward this type of training. Among the Swedish students, a follow-up was carried out after 6 months. Depending on the distributions, *t* tests or Mann-Whitney tests were used. Correlation between variables was assessed by using Spearman rank correlation. Regression analyses were used for time-dependent variables.

Results: The participants enjoyed the training and reported a self-perceived benefit as a consequence of training. The mean rating for self-efficacy increased from 5.8/7 (SD 0.72) to 6.5/7 (SD 0.57, $P<.001$). In the Swedish follow-up, it subsequently increased from 5.7/7 (SD 0.56) to 6.3/7 (SD 0.38, $P=.006$). In the Swedish group, the mean concentration value increased from 52.4/100 (SD 9.8) to 62.7/100 (SD 8.9, $P=.05$); in the US group, the concentration value increased from 70.8/100 (SD 7.9) to 82.5/100 (SD 4.7, $P<.001$). We found a significant positive correlation ($P<.001$) between self-efficacy and concentration scores. Overall, the participants were moderately or highly immersed and the software was easy to use.

Conclusions: By using online MMVWs, team training in CPR is feasible and reliable for this international group of high school students (Sweden and United States). A high level of appreciation was reported among these adolescents and their self-efficacy increased significantly. The described training is a novel and interesting way to learn CPR teamwork, and in the future could be combined with psychomotor skills training.

(J Med Internet Res 2013;15(1):e9) doi:10.2196/jmir.1715

KEYWORDS

Serious games; virtual learning environments; MMVW; avatars; students; cardiopulmonary resuscitation; patient simulation; self-efficacy; concentration

Introduction

Sudden cardiac arrest (SCA) is one of the most common causes of death [1]. Society depends heavily on laypersons in the initial resuscitation of cardiac arrest victims. To disseminate the necessary skills, training in cardiopulmonary resuscitation (CPR) is mandatory in many school systems. However, the effectiveness of standard CPR training among nonprofessionals [2-4] and health care professionals [5] has been questioned, mainly because of poor retention of skills and low self-efficacy among the trainees. Further, current CPR programs are focused on training of the individual. If several rescuers are present, lack of team coordination could hamper the effectiveness of CPR [6,7].

There is a demand for methods that are easy to deliver, require little material resources, and are highly engaging. Although a matter of both guideline simplicity and training methods, long-term retention is one ideal target for training. In search of new and more effective CPR training solutions, several alternative methods of delivery have been suggested during the past decade [8-12]. Screen- or simulator-based methods have been developed as alternatives to traditional manikin-based skills-oriented training. These alternatives primarily employ instructive video demonstrations or full-scale simulations.

Massively multiplayer virtual world (MMVW) technology has its roots in computer games and videogames for entertainment. Since the 1990s, videogames have been used for educational purposes in various fields, including medicine. These applications have been termed *serious games* [13]. For CPR training purposes, the option of virtual worlds for training has not been studied so far. Some encouraging results from a pilot study have demonstrated the potential of using virtual worlds for initial training of trauma team leaders [14].

In a recent study, we found that a MMVW serious game for CPR training was feasible for use among medical students. Although no evidence of stimulated recall of CPR procedures was found, the participants were enthusiastic and reported increased concentration during the training. Further, the participants' self-efficacy increased after training [15]. *Individually perceived efficacy* refers to beliefs about one's capabilities to learn or perform behaviors at designated levels [16]. Salas and Burke [17] have argued that for effective training, the individual characteristic of trainees, of which self-efficacy is one, should be further elucidated. Self-efficacy has been emphasized as an important feature of medical simulator training [18]. Brusso et al [19] have also pointed out the importance of self-efficacy during videogame-based training.

Information and computer technology (ICT) is becoming increasingly integrated into the everyday life of the younger generations. It has been suggested that this has implications on learning [20-23]. Although new methods for teaching based on ICT are being implemented for children and adolescents in school, knowledge is limited concerning effectiveness of such training in the medical field [24]. Taking advantage of the ICT skills of today's adolescents regarding serious games poses potential challenges [25-27]. It has been discussed that ICT-based teaching methods often face cultural challenges when implemented in a non-native context or environment [28,29]. This fact must not be ignored when initiating multinational training programs [30].

By using a deliberate practice model [31], the aim of this prospective international study (in the United States and in Sweden) was to examine the feasibility of using a MMVW for training high school students to respond appropriately to a medical emergency requiring CPR teamwork, and to examine their experiences, self-efficacy, and concentration. Comparing the results would indicate if these results are generalizable to the 2 different student populations.

Methods

Recruitment and Sample

The study was performed after institutional review board approval at both study locations. After obtaining informed consent, 12 high school students from the natural science program at Huddinge High School in Huddinge, Sweden, and 24 high school students from Woodside High School in Woodside, CA, USA, were recruited and enrolled on a voluntary basis. Recruitment was carried out by their teachers and through solicitation by announcements regarding the study. Before the practice, all students had participated in compulsory conventional CPR training at their respective schools within the past 6 months. Demographic data of all study participants are displayed in Table 1. There was no clear difference between groups in mean frequency of videogame play, but a higher percentage of low-frequency and high-frequency players were found in the Swedish group.

Table 1. Demographics of participating high school students from Sweden and the United States (N=36).

Characteristics		Country, n (%)	
		Sweden n=12	United States n=24
Sex			
	Female	5 (42)	15 (62)
	Male	7 (58)	9 (38)
Grade			
	10	12 (100)	23 (96)
	11	0	1 (4)
Frequency of videogame play			
	Less than once a month	5 (42)	5 (21)
	Every second week	2 (17)	11 (46)
	Once a week	1 (8)	5 (21)
	Several times every week	2 (17)	2 (8)
	Every day	2 (17)	1 (4)

Intervention and Curriculum Content

The study was collaboratively planned and implemented, although it was not implemented at exactly the same time. From previous work, a virtual-world platform was jointly developed with prehospital CPR team training capabilities (Forterra Systems Inc, San Mateo, CA, USA; OLIVE game development platform now the property of SAIC, McLean, VA, USA). The virtual-world environment was tailored according to real-world examples. In the virtual world, each participant was represented by an avatar that was controlled by the computer mouse and keyboard. The participants could communicate with one another in real time by means of a microphone and headset (Figures 1 and 2).

An overview of the protocol is shown in Figure 3. In the beginning of the training session, the participants rehearsed and updated their CPR knowledge. In the Swedish group, this was done by a short (10-minute) lecture, whereas the US group rehearsed the CPR procedure in the virtual environment. Familiarization to the virtual world and training in maneuvering and communicating within it was done just before the CPR training started.

In the CPR scenarios, the participants trained in mixed groups of 3 students. In the Swedish study group, CPR was practiced in 4 scenarios, in which the first was regarded as a familiarization scenario. In the US group, familiarization to the virtual world included guided practice and CPR was trained in 3 scenarios. The overall time for the training sessions (including data collection) was 90 to 120 minutes.

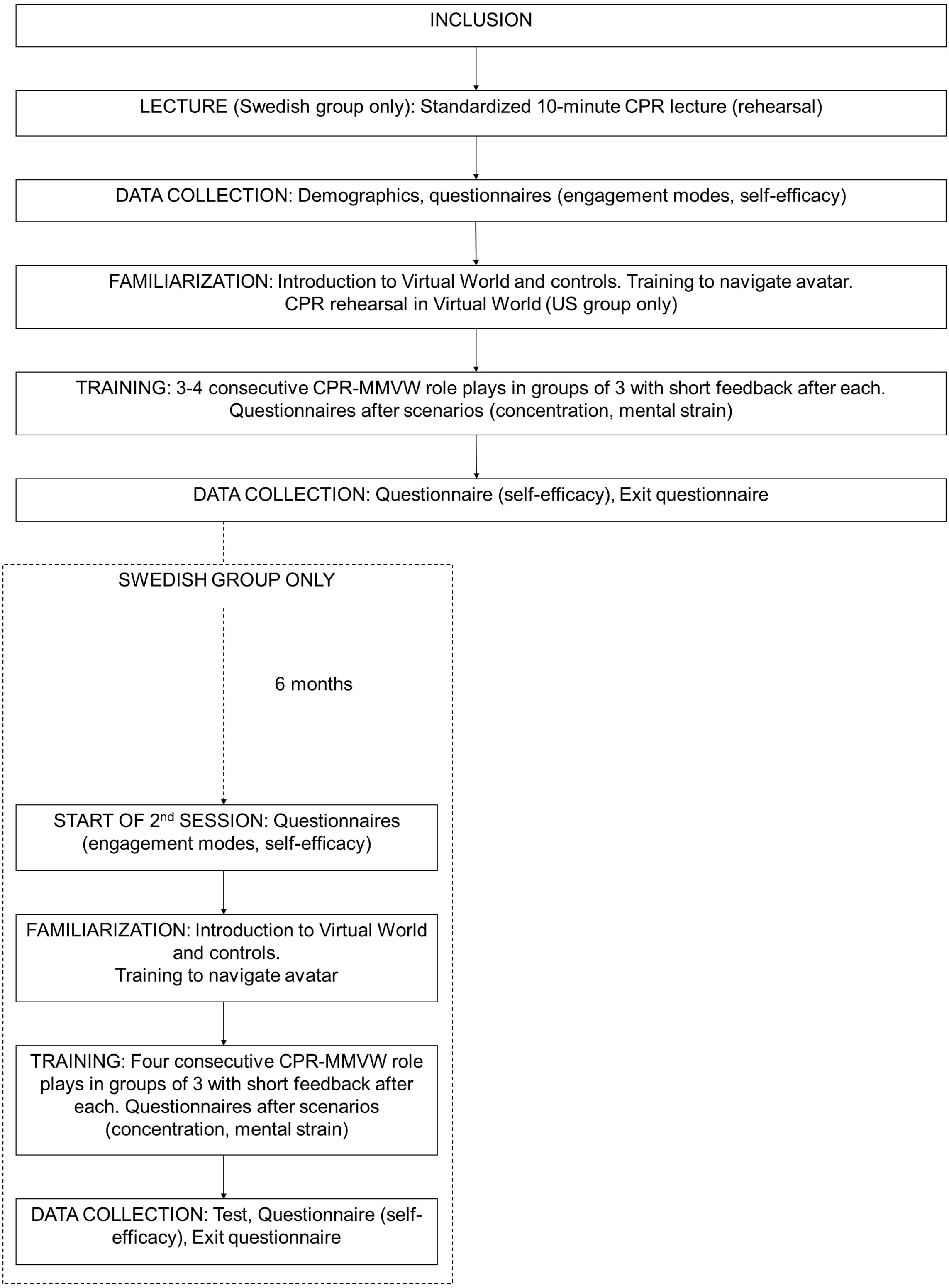
The trainees were instructed to approach a victim (an avatar that was controlled by an instructor) whose collapse they had witnessed in the virtual world, take the correct diagnostic steps, and collaboratively perform the cognitive and procedural measures associated with basic adult life support in accordance to the 2005 guidelines from the American Heart Association [32]. These included (1) moving to the victim, (2) checking the victim for consciousness, (3) declaring the victim unconscious, (4) checking the victim’s airway and breathing, (5) calling for help, (6) performing chest compressions and rescue breaths as stated by the CPR protocol, (7) relieving the rescuers, and (8) assisting the arriving paramedic and giving a brief report to him. The settings for the virtual-world scenarios were in a classroom and in an outdoor parking lot. After each scenario, an instructor gave short feedback about their CPR performance compared to the guidelines.

Initially a follow-up was planned in both locations, but it had to be abandoned in the United States because of difficulties reassembling the study group. In the Swedish study group, a prospective test-retest explorative design could be used. This included a second session 6 months (177-203 days) after the first training. Based on the literature, 6 months was chosen as the time between the training sessions [33]. The second session was identical to the first except that the lecture on CPR was excluded. No participants in the Swedish group dropped out between these 2 training sessions.

Figure 1. Screenshot of avatar performing chest compressions on a victim in the virtual world (parking lot scenario).

Figure 2. Screenshot of avatar performing chest compressions on a victim in the virtual world (classroom scenario) while talking to relieving paramedic.

Figure 3. Design of the study.



Instruments and Measurements

Self-efficacy, a construct resting on Bandura’s social cognitive theory [16], has been found to be a strong predictor of actual performance. Because we recently found that medical students’ self-efficacy increased after training CPR in a MMVW serious game [15], we examined if these results could be replicated in a group of young laypersons (ie, high school students). Self-efficacy was self-assessed before and after each training session by using a 5-item questionnaire in which each item was rated on a 7-grade Likert-type scale. The value for self-efficacy was calculated as the mean of these items as described elsewhere [34].

In the same study [15], medical students' concentration increased during training. Concentration is theoretically well connected to the concept of intrinsic motivation in performing a task. A higher level of concentration is indicative of being captured by an activity and a greater probability of continuing with it. In the present study, we assessed the high school students' concentration as an index of their motivation in the training tasks. The greater the self-efficacy, the more active the efforts would be [16]. For assessing concentration we used 8 items from already validated instruments [35,36]. In the Swedish subgroup, concentration was self-assessed after the first, second, and fourth scenario during both sessions. In the US study group, concentration was self-assessed after the first and third scenarios.

All individuals' perceptions regarding the feasibility of training were assessed in an exit questionnaire that was distributed after the training sessions. This included the level of perceived immersion in the virtual world, questions covering technical aspects (ie, technical difficulties, ease of use), and usefulness of the training. These assessments were made by using 5-grade Likert-type scales. Control questions about self-confidence were given and the Swedish participants were also asked if this training mode could have a role in future education; 5-grade Likert-type scales were used for both questions.

In the exit questionnaire, the Swedish participants were asked to specifically comment on the perceived strengths and weaknesses of the scenarios. Meaning-bearing units [37] were identified in these comments and consequently analyzed, classified, and categorized by one of the authors (JC). If the same answer was given several times by a single participant, it was counted as only 1 occurrence to have it justly weighted during analysis. All participants were asked if and how the training had changed their feelings/perceptions about having to act in an emergency situation. In these comments, meaning-bearing entities were also identified, analyzed, classified, and categorized by the one of the authors (JC).

Although self-efficacy beliefs exercise a powerful influence on human action [16], many other psychological processes can affect the strength of this relationship. We assumed that engagement modes might potentially affect the strength of relationship between self-efficacy beliefs and behavior differently among participants. Engagement modes are different ways of interacting with new information technologies [35] and, in this study, students' different thoughts, feelings, and purposes they had toward CPR training using avatars in MMVM. We calculated an index of the negative engagement modes (sum of scores for the frustration/anxiety and hesitation/avoidance modes), as described by Hedman and Sharafi [36], because they may be associated with lower self-efficacy. Negative engagement modes were assessed before each session from a 15-item questionnaire using a 5-point Likert-type scale.

We assumed that very high mental strain could be another factor that may jeopardize perceived self-efficacy. Only in the Swedish subgroup, mental strain could be self-assessed after the first, second, and fourth scenario during each session. Mental strain was measured using Borgs' CR10 scale (0 =no mental strain at all; 10 = extremely high mental strain) after each session [38].

Data Analysis

Participants with single missing data were excluded from that particular analysis. Repeated measurements analysis was used to analyze time-dependent data (concentration) and regression analysis was used to evaluate the dependency between these variables. For correlation between concentration and self-efficacy, Spearman rank correlation (ρ) was performed. Statistical comparisons were made by using the Mann-Whitney test or the Student *t* test, after validation for normal distribution by using the Shapiro-Wilk test. The significance level was set at *P*<.05. The calculations were performed by using SigmaStat version 3.5 (Systat Software Inc, Point Richmond, CA, USA). Data are presented as mean and standard deviation (SD) or median and interquartile range (IQR) depending on the type of distribution.

Results

Summative data from Sweden and the United States were analyzed as one group, except for concentration in which there were large differences between the groups.

Controlled Factors

Negative engagement modes in the Swedish study group occurred during session 1 (mean 5.4, SD 1.1) and during session 2 (mean 5.4, SD 1.1). In the US study group, the mean was 3.4 (SD 0.3). The difference between the groups was significant (*P*<.001). The mean value of mental strain, measured only in the Swedish group, was low (mean 3.4, SD 1.7) without significant change over time.

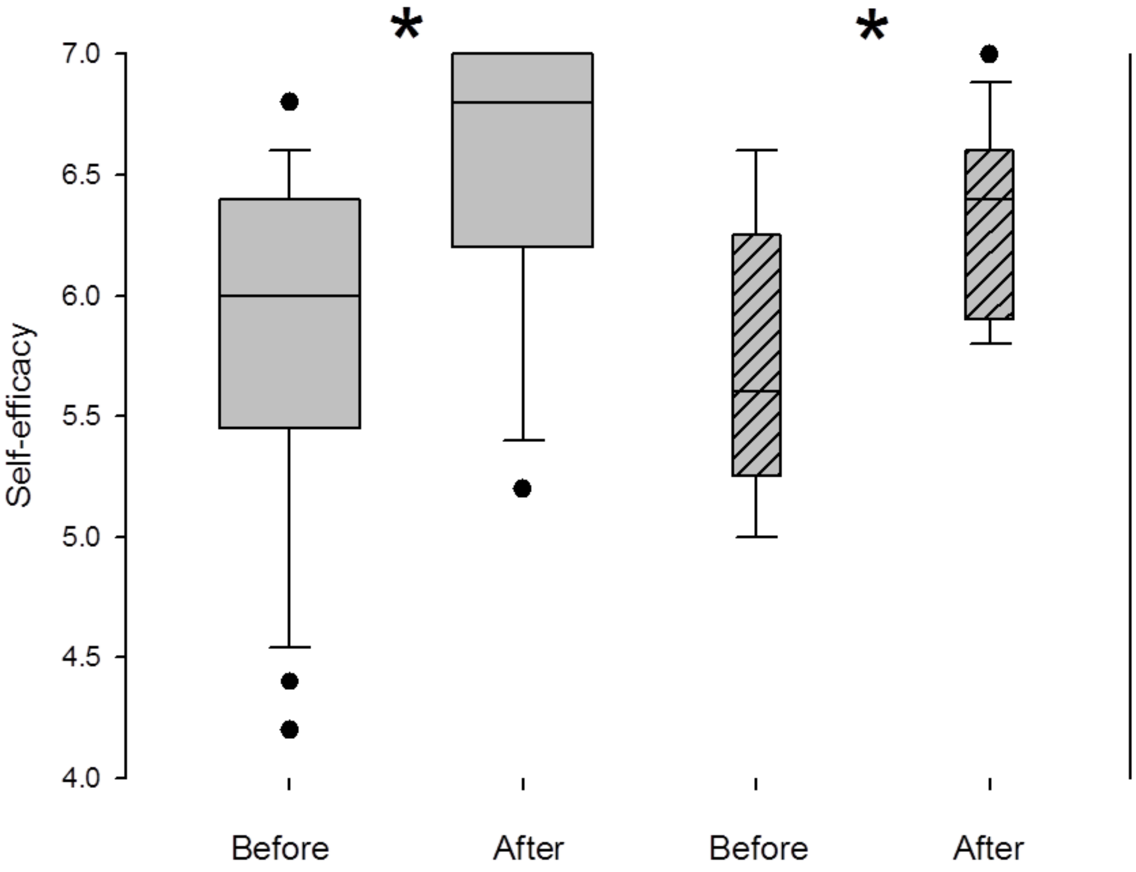
Self-efficacy and Concentration

Self-efficacy increased significantly after training compared to before training (Figure 4). In the Swedish group, this was also replicated during the second training session.

In general, the concentration (task motivation) level was in the medium to high range. However, the results in the US group were significantly higher than the Swedish group (*P*<.001); in the latter, the mean concentration score showed a significant increase over time (scenario 1 of session 1: mean 52.4, SD 9.8; scenario 4 of session 2: mean 62.7, SD 8.9; *P*<.05). In the US group, the mean concentration score increased significantly from 70.8 (SD 7.9) to 82.5 (SD 4.7, *P*<.001).

Self-efficacy was positively related to concentration during training before and after training (self-efficacy before training and concentration during first scenario: ρ = 0.49, *P*=.006; self-efficacy after training and concentration during last scenario: ρ = 0.60, *P*<.001).

Figure 4. Self-efficacy in the study group. The two left-most boxes present self-efficacy before and after training (N = 36).The two right-most boxes (striped) refer to the second session during the Swedish part of the study (n = 12). The box-blot illustrates 25th and 75th percentiles with median value as a solid line inside the plot and whiskers showing 10th and 90th percentiles (outliers marked outside this). Significance (P <.05) between the measurements is denoted with an asterisk (*).



Perceptions of the Training and Attitudes

The exit questionnaire contained information about the participants’ experiences of the technical aspects of the virtual world and their opinions about the MMVW CPR team training. The participants’ ratings indicated a significant change of confidence after the training (*P*<.001). Numerical ratings are summarized in Table 2. Table 3 summarizes perceived strengths and weaknesses with the simulated scenarios, investigated in the Swedish subgroup.

Twenty-eight of the 36 participants (78%) agreed that the MMVW experience had changed their feelings or perceptions about responding to a medical emergency, without any differences between study groups (Table 4).

Table 2. Exit questionnaire about the participants' perceptions of the training and attitudes toward it based on 5-point Likert-type scales.

Exit question	Swedish group, median (IQR)[a]		US group, median (IQR)[a]
	Session 1 n=12	Session 2 n=12	n=24
Did you feel that you were actually there?[b]	4.0 (3.5-4.5)	4.0 (3.0-4.0)	4.0 (4.0-4.0)
Did you experience any technical difficulties?[b]	2.0 (1.5-2.5)	2.0 (1.0-2.5)	2.0 (2.0-3.0)
How easy to learn to control your avatar?[c]	4.0 (3.5-4.5)	4.0 (3.5-5.0)	4.0 (4.0-5.0)
How useful for learning to react to a medical emergency?[d]	—	4.5 (4.0-5.0)	4.0 (4.0-5.0)
Do you think this type of simulated training has a part in the education of tomorrow?[e]	5.0 (4.0-5.0)	5.0 (5.0-5.0)	—
How confident to react to a medical emergency before today's session?[f]	—	4.0 (3.0-4.0)	2.5 (2.0-3.0)
How confident to react to a medical emergency after today's session?[f]	—	4.2 (4.0-5.0)	4.0 (4.0-4.5)

[a] IQR: interquartile range; —: question was not asked.

[b] 1 = not at all; 5 = all of the time.

[c] 1 = never learned how; 5 = very easy.

[d] 1 = not useful; 5 = very useful.

[e] 1 = not at all; 5 = yes, absolutely.

[f] 1 = not confident; 5 = extremely confident.

Table 3. The Swedish participants' (n=12) answers about strengths and weaknesses of the simulated scenarios.

Category		n (%)
Strengths		
	Suitable and realistic environment	9 (28)
	Good way to repeatedly practice and learn	8 (25)
	Necessary to adapt to changing circumstances	2 (6)
	Training teamwork aspects	7 (22)
	Good in general	6 (19)
Weaknesses		
	Too easy tasks, more options wanted	10 (30)
	Lack of realism and a richer environment	8 (24)
	Technical problems	9 (27)
	Requires familiarization	6 (18)

Table 4. Statements about how the training changed the participants' (n=36) feelings or perceptions about responding to a medical emergency (participants could give several answers).

Category	n (%)
Work better in a CPR team	11 (26)
Increased confidence for such an emergency situation	20 (47)
Better CPR knowledge	12 (28)

Discussion

Our ambition in this feasibility study was to selectively analyze how teams of young laypersons, personified as avatars, reacted toward and interacted within a virtual world for teaching how to respond appropriately to a medical emergency requiring CPR. Our study sample were high school students in an international setting. In this study, we refined and used a virtual world for CPR team training previously studied in medical students [15]. The high school students clearly appreciated the way the training could influence real-world behaviors in several aspects and it was associated with an increased self-efficacy in both countries. Our results are consistent with previous findings in a group of medical students [15]. These findings suggest that this completely novel method for CPR training is also feasible and reliable in young laypersons.

The increase in students' self-efficacy has 2 important implications. First, it illustrates the beliefs of the trainee (ie, that the level of control of required skills increased). We triangulated this finding with the control questions in the exit questionnaire about the participants' self-confidence that also increased because of the training. Secondly, which may be particularly important concerning CPR teamwork, it gives an indication that the trainee actually would feel more prepared to act in a real-world CPR event. This assumption should be tested in future studies.

Our finding that there was a positive correlation between self-efficacy and concentration (as an index of motivation in the training tasks) is in accordance with Bandura's [16] social cognitive theory in which self-efficacy beliefs influence cognitive, motivational, affective, and decisional processes. The course-effect relationship between self-efficacy and concentration should, however, be elucidated by using a control group design. Concentration and the perception of immersion, characteristics of computer games, are of great importance because these are signs of how the trainees are engaged in the training. A high level of task involvement is a prerequisite for active learning. Concentration has been positively associated with effective teaching and learning [39]. Current CPR training seems to carry problems in these aspects, which might lead to a lower degree of influence on the trainees and, in turn, less retention.

An important aspect is the degree of realism and sense of presence in virtual-world training. There certainly is a risk that the participants may perceive the training environment as awkward and experience a lack of real-world resemblance. In the past few years, virtual-world training has been used extensively for cybertherapy in psychiatry in which these aspects are crucial. Studies from this field show that this type of training has potential for immersion and active engagement [40,41]. Also, our results indicate that the learning scenarios and technology developed for our study were effective for most students in achieving a “suspension of disbelief.”

Although the need for CPR training is universal, computer habits as well as pedagogical approaches and attitudes toward e-learning differ greatly among different countries and cultures [42,43]. Attitudes toward team interactions in virtual settings seem to be of great importance and may also differ. Ahanchian and McCormick [28] created a framework based on Hofstede's cultural dimensional framework to assist understanding of cultural differences. With that in mind, we expanded the study to another contextual and cultural setting to analyze the generalizability of MMVW as a new teaching method for team-based CPR training. Videogame habits among the participants in the 2 different countries were somewhat different. In particular, the Swedish group contained a larger group of participants with less computer gaming activity. Although there is some variation in the literature, videogame habits among the participants did not seem to be above average compared to larger US samples [44,45]. Wong [29] has pointed out potential difficulties with e-learning in multicultural settings, in which cultural influences may impact how the learning is perceived and its effectiveness. These problems primarily have their roots in cultural-sensitive issues, such as attitudes toward how one should behave in an e-learning situation or how active and self-directed the learner is expected to be. In the present MMVW, the only interaction between participants was by direct voice interaction in the virtual world, and the interaction with the teacher or instructor was either in-world (by the use of avatars) or in the real classroom. Hence, the virtual world itself will not be affected by cultural issues, except for appearances of objects (eg, buildings, avatars) and the text on the interactive buttons used to examine and treat the victim. On the other hand, Uzuner [46] pointed out in a review that attitudes toward distance learning may differ greatly among different cultures, in particular between so-called collectivistic and individualistic cultures. Without further data from our international study, we cannot draw any conclusion about how culture influences our findings. However, the aim was to study if our findings would be similar in the 2 countries and, hence, more generalizable. As with leisure gaming MMVW, our data quite clearly indicate that users in different countries enjoy similar engagement in this MMVW CPR team training.

The major difference in the individual experiences during training between the 2 countries was in concentration, where the US sample was on a considerably higher level, in general. In this group, the negative engagement modes in how one relates to information technology was also considerably lower. This difference in attitude toward the training method, indicating a more relaxed and open attitude toward ICT in the US sample and reflecting the smaller percentage of participants with low-frequency computer gaming activity, may explain some of the difference in concentration. Another reasonable explanation could be that the educational software was more developed and stable, and technical and educational support was more easily accessible at that time. During the training in the United States, developers and technical expertise was readily present.

The individual comments indicate that MMVW team CPR training was appreciated for its team aspects and for being easy to practice repetitively leading to better knowledge and confidence. On the other hand, among the Swedish participants, remarks were made about the tasks being too easy and with too few options for interaction. Such indications warrant further investigation. One of the appealing features of serious games are their ability to vary the level of difficulty and keep the user in a high level of engagement—these comments indicate that more could be accomplished in these aspects, possibly leading to higher levels of concentration.

The results from these 2 groups are in accordance with what we have found when studying a group of medical students [15]. This implies the feasibility of the concept that MMVW for training CPR teamwork can be carried out easily and with good acceptance from the trainees. Personal observations, reviews of ratings, and verbal and written comments indicate that this new learning technology is highly appreciated by high school participants. Although CPR was the training content for this study, virtual worlds may be used for other medical training purposes. Patel and coworkers [47] have used an interactive virtual-world session for introducing the operating room environment to novices. Compared to standard real-world introduction, they found the virtual-world introduction to be as beneficial for gaining knowledge, skills, and attitudes. However when it comes to training of team activities, there are few reports in the medical literature.

Important issues when designing CPR training methods are cost and adaptability [48,49]. Not only is it important to find an engaging and popular form of training, it must also be associated with reasonable costs and carry the advantages of being easily adopted to different target groups as well as future changes in CPR guidelines. Further, it is also of great importance that the system and method used is easy to operate from a teacher's perspective. A complex method requiring high computer interest and skills from the teacher's side can be expected to be difficult to implement.

In society, initiation of action in sudden emergency medical situations depends on the vigilance of bystanders. The problem of delivering up-to-date education and training in this field for all citizens is huge. To deal with retention problems through recurrent training is perhaps even more challenging. Although some evidence of good retention after conventional training in school-aged children exists [50], many reports are discouraging.

Today, manikin-based CPR training is the norm. Such training carries several weaknesses (eg, accessibility, cost, engagement among trainees, lack of team aspects, and low long-term efficacy). Some of these issues may be addressed by the use of MMVW CPR team training. Further, the uniqueness of this training not only lies in the different medium for interaction and transmission, but also in the possibility to train CPR in a "live" setting where the medical emergency is put into context. Also, flexibility in changing the virtual environment, number of rescuers, and level of difficulty and complexity is increased compared to standard CPR training. On the other hand, using virtual worlds for training creates new demands, such as computer hardware and Internet connections, trainee computer skills, and time for learning the software and familiarizing to the virtual environment.

Current pedagogical training concepts advocate the division of complex tasks into the training of its included parts during basic training [51]. CPR consists of cognitive as well as team and psychomotor skills. Hence, it might be beneficial to learn and train the different aspects of CPR in different modes and at different times. When the separate parts are mastered, they may be integrated and trained. In its current format, MMVW CPR only consists of some parts of the whole, cognitive, and team aspects, whereas the psychomotor skills are not covered. To cover the full extent of bystander CPR (or other forms of CPR), this training modality has to be developed further.

The main goal for this study was not to validate or test a particular virtual world, but rather to explore the concept and study feasibility aspects. Although manikin-based training continues to be the prevailing method for learning the basic psychomotor skills of CPR, this study demonstrates the potential added value of MMVW for situated and team-based learning in which laypersons are able to practice the sequence of actions necessary to respond appropriately to different medical emergencies. One of the most obvious advantages with virtual-world training is its almost endless variability. Changes in virtual environment, equipment, and situation creates potential means for practicing a wide variety of medical tasks, many that normally would be difficult to practice in the real world.

A limitation of this work is the sample size. The originally planned follow-up in the United States was abandoned because of a dropout rate of almost 50%. This occurred because the follow-up had to be scheduled in the subsequent school year. Students had either left the district, changed schools, or were in classes in which their absence was not allowed which precluded participation in the study follow-up. It certainly turns the focus to the difficulties in carrying out multinational field studies in which adherence to initial protocol was restrained because of cultural and curriculum circumstances. In future studies, this issue has to be carefully addressed in close cooperation with the school administrators and staff. Our intention was to study the samples at a time when CPR training was part of their normal curriculum. Further, the study was carried out in schools situated in 2 high-income Western countries. Although we found the same patterns in the results in the 2 countries, these results would be easier to generalize if more groups were studied (ie, different cultures and ages). Future studies should investigate retention after training and transfer to real-world CPR. Also, more data on the teachers' perceptions are warranted.

Conclusions

The data from this feasibility study support the use of MMVWs for teaching high school students to respond appropriately to a medical emergency requiring CPR teamwork. It was feasible, reliable, and enjoyed by digital natives in Sweden and in the United States. This study demonstrates the potential added value of MMVW for situated learning in which young laypersons are able to practice the sequence of actions necessary to respond appropriately to different medical emergencies.

Even with a great demand for new training methods in high school and in the area of CPR training, MMVW serious games must be carefully studied not only from the students' perspective, but also from the teachers' and organizations' perspectives to clarify the challenges and needs required for implementation.

Acknowledgments

We would like to express our gratitude to Parvati Dev formerly at SUMMIT, Stanford; Carl-Johan Wallin at the Center for Advanced Medical Simulation and Training; Christopher Medin, Department of Biomedical Technology at Karolinska University

Hospital; Karl Stengård, Huddinge High School; and Arnold Hendrick and Laura Kusumoto, previously at Forterra Inc, San Mateo, CA, for valuable contributions in planning of the study as well as development and support of the software. Further, we would like to express our gratitude to Laura Peridikomatis at Woodside High School, Woodside, CA, for making these studies possible. Also, we thank Per Näsman at the Royal Institute of Technology, Stockholm, for assistance with statistical methods and analyses. This study was fully financed by research grants from the Wallenberg Global Learning Network (WGLN), Stockholm County Council and EU Research Program Goal 1, Northern Sweden.

Conflicts of Interest

WLH and LFT were the principal investigators on the Wallenberg grant that funded Forterra Systems's development of the virtual environments used in this study. The grant also paid WLH and Patricia Youngblood a small portion of their salaries at SUMMIT Stanford.

References

1. American Heart Association. 2010. CPR & Sudden Cardiac Arrest (SCA) Fact Sheet URL: http://www.heart.org/HEARTORG/CPRAndECC/WhatisCPR/CPRFactsandStats/CPR-Statistics_UCM_307542_Article.jsp [accessed 2010-12-14] [WebCite Cache ID 5uyfqHQ5O]
2. Kaye W, Mancini ME. Teaching adult resuscitation in the United States--time for a rethink. Resuscitation 1998 Jun;37(3):177-187. [Medline: 9715778]
3. Lim GH, Seow E. Resuscitation for patients with out-of-hospital cardiac arrest: Singapore. Prehosp Disaster Med 2002;17(2):96-101. [Medline: 12500733]
4. Swor R, Khan I, Domeier R, Honeycutt L, Chu K, Compton S. CPR training and CPR performance: do CPR-trained bystanders perform CPR? Acad Emerg Med 2006 Jun;13(6):596-601. [doi: 10.1197/j.aem.2005.12.021] [Medline: 16614455]
5. Wik L, Kramer-Johansen J, Myklebust H, Sørebø H, Svensson L, Fellows B, et al. Quality of cardiopulmonary resuscitation during out-of-hospital cardiac arrest. JAMA 2005 Jan 19;293(3):299-304. [doi: 10.1001/jama.293.3.299] [Medline: 15657322]
6. Cooper S, Wakelam A. Leadership of resuscitation teams: "Lighthouse Leadership". Resuscitation 1999 Sep;42(1):27-45. [Medline: 10524729]
7. Marsch SC, Müller C, Marquardt K, Conrad G, Tschan F, Hunziker PR. Human factors affect the quality of cardiopulmonary resuscitation in simulated cardiac arrests. Resuscitation 2004 Jan;60(1):51-56. [doi: 10.1016/j.resuscitation.2003.08.004] [Medline: 14987784]
8. 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]
9. Jones I, Handley AJ, Whitfield R, Newcombe R, Chamberlain D. A preliminary feasibility study of a short DVD-based distance-learning package for basic life support. Resuscitation 2007 Nov;75(2):350-356. [doi: 10.1016/j.resuscitation.2007.04.030] [Medline: 17618033]
10. Toner P, Connolly M, Laverty L, McGrath P, Connolly D, McCluskey DR. Teaching basic life support to school children using medical students and teachers in a 'peer-training' model--results of the 'ABC for life' programme. Resuscitation 2007 Oct;75(1):169-175. [doi: 10.1016/j.resuscitation.2007.03.009] [Medline: 17482334]
11. Wayne DB, Didwania A, Feinglass J, Fudala MJ, Barsuk JH, McGaghie WC. Simulation-based education improves quality of care during cardiac arrest team responses at an academic teaching hospital: a case-control study. Chest 2008 Jan;133(1):56-61. [doi: 10.1378/chest.07-0131] [Medline: 17573509]
12. Wik L, Myklebust H, Auestad BH, Steen PA. Retention of basic life support skills 6 months after training with an automated voice advisory manikin system without instructor involvement. Resuscitation 2002 Mar;52(3):273-279. [Medline: 11886733]
13. Westera W, Nadolski RJ, Hummel HGK, Wopereis IGJH. Serious games for higher education: a framework for reducing design complexity. J Comput Assist Learn 2008;24(5):420-432 [FREE Full text] [doi: doi: 10.1111/j.1365-2729.2008.00279.x]
14. Youngblood P, Harter PM, Srivastava S, Moffett S, Heinrichs WL, Dev P. Design, development, and evaluation of an online virtual emergency department for training trauma teams. Simul Healthc 2008;3(3):146-153. [doi: 10.1097/SIH.0b013e31817bedf7] [Medline: 19088658]
15. Creutzfeldt J, Hedman L, Medin C, Heinrichs WL, Felländer-Tsai L. Exploring virtual worlds for scenario-based repeated team training of cardiopulmonary resuscitation in medical students. J Med Internet Res 2010;12(3):e38 [FREE Full text] [doi: 10.2196/jmir.1426] [Medline: 20813717]
16. Bandura A. Self-Efficacy: The Exercise of Control. New York: WH Freeman; 1997.
17. Salas E, Burke CS. Simulation for training is effective when. Qual Saf Health Care 2002 Jun;11(2):119-120 [FREE Full text] [Medline: 12448801]
18. Kneebone R. Evaluating clinical simulations for learning procedural skills: a theory-based approach. Acad Med 2005 Jun;80(6):549-553. [Medline: 15917357]
19. Brusso RC, Orvis KA, Bauer KN, Tekleab AG. Interaction among self-efficacy, goal orientation, and unrealistic goal-setting on videogame-based training performance. Military Psychology 2012;24(1):1-18. [doi: 10.1080/08995605.2012.639669]

20. Carr D. London Knowledge Lab, Institute of Education, University of London. 2007 May. Digital technologies, learning and game formats: Computer games, motivation and gender in educational contexts URL: http://playhouse.files.wordpress.com/2007/04/editfinalreportdigitalgames.pdf [accessed 2010-12-14] [WebCite Cache ID 5uyiGPH6Q]
21. Oblinger DG. Journal of Interactive Media in Education. 2004. The next generation of educational engagement URL: http://jime.open.ac.uk/2004/8/oblinger-2004-8-disc-paper.html [accessed 2009-10-29] [WebCite Cache ID 5ktca0OHO]
22. Prensky M. Digital natives, digital immigrants. On the Horizon 2001;9(5) [FREE Full text]
23. Squire K. Changing the game: what happens when video games enter the classroom? Innovate 2005;1(6) [FREE Full text]
24. Bond WF, Lammers RL, Spillane LL, Smith-Coggins R, Fernandez R, Reznek MA, Society for Academic Emergency Medicine Simulation Task Force. The use of simulation in emergency medicine: a research agenda. Acad Emerg Med 2007 Apr;14(4):353-363. [doi: 10.1197/j.aem.2006.11.021] [Medline: 17303646]
25. Broos A. Gender and information and communication technologies (ICT) anxiety: male self-assurance and female hesitation. Cyberpsychol Behav 2005 Feb;8(1):21-31. [doi: 10.1089/cpb.2005.8.21] [Medline: 15738690]
26. Evans N, Mulvihill TM, Brooks NJ. Mediating the tensions of online learning with Second Life. Innovate 2008;4(6) [FREE Full text]
27. Bryce J, Rutter J. Gender dynamics and the social and spatial organization of computer gaming. Leisure Studies 2003;22(1):1-15. [doi: 10.1080/02614360306571]
28. Ahanchian MR, McCormick J. Culture and the processes of virtual teaming for training. J Comp Assist Learn 2009;25(4):386-396 [FREE Full text] [doi: 10.1111/j.1365-2729.00314.x]
29. Wong AL. Cross-cultural delivery of e-learning programmes: Perspectives from Hong Kong. International Review of Research in Open and Distance Learning 2007;8(3) [FREE Full text]
30. Sadykova G, Dautermann J. Crossing cultures and borders in international online distance higher education. Journal of Asynchronous Learning Networks 2009;13(2):89-114 [FREE Full text]
31. Ericsson KA. Deliberate practice and acquisition of expert performance: a general overview. Acad Emerg Med 2008 Nov;15(11):988-994. [doi: 10.1111/j.1553-2712.2008.00227.x] [Medline: 18778378]
32. ECC Committee, Subcommittees and Task Forces of the American Heart Association. 2005 American Heart Association Guidelines for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care. Circulation 2005 Dec 13;112(24 Suppl):IV1-203 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.105.166550] [Medline: 16314375]
33. Mäkinen M, Niemi-Murola L, Mäkelä M, Castren M. Methods of assessing cardiopulmonary resuscitation skills: a systematic review. Eur J Emerg Med 2007 Apr;14(2):108-114. [doi: 10.1097/MEJ.0b013e328013dc02] [Medline: 17496690]
34. Pintrich PR, Smith DA, Garcia T, McKeachie WJ. Reliability and predictive validity of the motivated strategies for Learning Questionnaire (MSLQ). Educ Psychol Meas 1993;53(5):801-813. [doi: 10.1177/0013164493053003024]
35. Csikszentmihalyi M, Abuhamedeh S, Nakamura J. Flow. In: Elliot A, Dweck CS, editors. Handbook of Competence and Motivation. New York: The Guilford Press; 2005:598-608.
36. Hedman L, Sharafi P. Early use of internet-based educational resources: effects on students' engagement modes and flow experience. Behaviour & Information Technology 2004;23(2):137-146 [FREE Full text] [doi: 10.1080/01449290310001648251]
37. Malterud K. Qualitative Methods in Medical Research [in Swedish]. Lund, Sweden: Studentlitteratur; 2009.
38. Borg G, Borg E. A new generation of scaling methods: Level-anchored ratio scaling. Psychologica 2001;28:15-45 [FREE Full text]
39. Csikszentmihalyi M, Rathunde K, Whalen S. Talented teenagers: the roots of success and failure. Cambridge, UK: Cambridge University Press; 1993.
40. Kenny PG, Parsons TD, Rothbaum B, Difede J, Reger G, Rizzo A. Optimizing clinical training for the treatment of PTSD using virtual patients. Stud Health Technol Inform 2009;144:264-268. [Medline: 19592778]
41. Walshe D, Lewis E, O'Sullivan K, Kim SI. Virtually driving: are the driving environments "real enough" for exposure therapy with accident victims? An explorative study. Cyberpsychol Behav 2005 Dec;8(6):532-537. [doi: 10.1089/cpb.2005.8.532] [Medline: 16332164]
42. Erumban AA, de Jong SB. Cross-country differences in ICT adoption: a consequence of culture? Journal of World Business 2006;41(4):302-314. [doi: 10.1016/j.jwb.2006.08.005]
43. Arenas-Gaitan J, Ramirez-Correa PE, Rondan-Cataluna FJ. Cross cultural analysis of the use and perceptions of web based learning systems. Computers & Education 2011;57:1762-1774. [doi: 10.1016/j.compedu.2011.03.016]
44. The Entertainment Software Association. 2010. Industry facts URL: http://www.theesa.com/facts/pdfs/ESA_Essential_Facts_2010.PDF [accessed 2011-05-24] [WebCite Cache ID 5yvZKYoCE]
45. Lenhart A, Kahne JK, Middaugh E, Rankin Macgill A, Evans C, Vitak J. Teens, video games, and civics. Washington, DC: Pew Internet & American Life Project; 2008 Sep 16. URL: http://www.pewinternet.org/~/media//Files/Reports/2008/PIP_Teens_Games_and_Civics_Report_FINAL.pdf.pdf [accessed 2011-05-24] [WebCite Cache ID 5yvaIYY2n]
46. Uzuner S. Questions of culture in distance learning: A research review. International Review of Research in Open and Distance Learning 2009;10(3) [FREE Full text]

47. Patel V, Aggarwal R, Osinibi E, Taylor D, Arora S, Darzi A. Operating room introduction for the novice. Am J Surg 2012 Feb;203(2):266-275. [doi: 10.1016/j.amjsurg.2011.03.003] [Medline: 21703594]
48. Hansen MM. Versatile, immersive, creative and dynamic virtual 3-D healthcare learning environments: a review of the literature. J Med Internet Res 2008;10(3):e26 [FREE Full text] [doi: 10.2196/jmir.1051] [Medline: 18762473]
49. Graafland M, Schraagen JM, Schijven MP. Systematic review of serious games for medical education and surgical skills training. Br J Surg 2012 Oct;99(10):1322-1330. [doi: 10.1002/bjs.8819] [Medline: 22961509]
50. Fleischhackl R, Nuernberger A, Sterz F, Schoenberg C, Urso T, Habart T, et al. School children sufficiently apply life supporting first aid: a prospective investigation. Crit Care 2009;13(4):R127 [FREE Full text] [doi: 10.1186/cc7984] [Medline: 19646229]
51. Kahol K, Vankipuram M, Smith ML. Cognitive simulators for medical education and training. J Biomed Inform 2009 Aug;42(4):593-604. [doi: 10.1016/j.jbi.2009.02.008] [Medline: 19269350]

Abbreviations

CPR: cardiopulmonary resuscitation
ICT: information and computer technology
IQR: interquartile range
MMVW: massively multiplayer virtual world
SCA: sudden cardiac arrest

Edited by G Eysenbach; submitted 18.12.10; peer-reviewed by AS Lu; comments to author 22.04.11; revised version received 06.08.12; accepted 03.10.12; published 14.01.13.

Please cite as:
Creutzfeldt J, Hedman L, Heinrichs L, Youngblood P, Felländer-Tsai L
Cardiopulmonary Resuscitation Training in High School Using Avatars in Virtual Worlds: An International Feasibility Study
J Med Internet Res 2013;15(1):e9
URL: http://www.jmir.org/2013/1/e9/
doi:10.2196/jmir.1715
PMID:23318253

Original Paper

Health Professional Learner Attitudes and Use of Digital Learning Resources

Stephen Maloney[1], B.Physio, MPH, PhD; Michael Chamberlain[1], B.Sci, B.Physio (Hons); Shane Morrison[2], B.Comp; George Kotsanas[2], GradDip Comp., PhD; Jennifer L Keating[2], GradDip (Manip.), PhD; Dragan Ilic[3], BSc, MSc, PhD

[1]Monash University, Department of Physiotherapy, Frankston, Australia

[2]Monash University, Faculty of Medicine, Nursing & Health Sciences, Clayton, Australia

[3]Monash University, Department of Epidemiology & Preventive Medicine, School of Public Health & Preventive Medicine, Monash University., Melbourne, Australia

Corresponding Author:
Stephen Maloney, B.Physio, MPH, PhD
Monash University
Department of Physiotherapy
Building B, McMahons Rd
Frankston, 3199
Australia
Phone: 61 399044240
Fax: 61 399044812
Email: stephen.maloney@monash.edu

Abstract

Background: Web-based digital repositories allow educational resources to be accessed efficiently and conveniently from diverse geographic locations, hold a variety of resource formats, enable interactive learning, and facilitate targeted access for the user. Unlike some other learning management systems (LMS), resources can be retrieved through search engines and meta-tagged labels, and content can be streamed, which is particularly useful for multimedia resources.

Objective: The aim of this study was to examine usage and user experiences of an online learning repository (Physeek) in a population of physiotherapy students. The secondary aim of this project was to examine how students prefer to access resources and which resources they find most helpful.

Methods: The following data were examined using an audit of the repository server: (1) number of online resources accessed per day in 2010, (2) number of each type of resource accessed, (3) number of resources accessed during business hours (9 am to 5 pm) and outside business hours (years 1-4), (4) session length of each log-on (years 1-4), and (5) video quality (bit rate) of each video accessed. An online questionnaire and 3 focus groups assessed student feedback and self-reported experiences of Physeek.

Results: Students preferred the support provided by Physeek to other sources of educational material primarily because of its efficiency. Peak usage commonly occurred at times of increased academic need (ie, examination times). Students perceived online repositories as a potential tool to support lifelong learning and health care delivery.

Conclusions: The results of this study indicate that today’s health professional students welcome the benefits of online learning resources because of their convenience and usability. This represents a transition away from traditional learning styles and toward technological learning support and may indicate a growing link between social immersions in Internet-based connections and learning styles. The true potential for Web-based resources to support student learning is as yet unknown.

(J Med Internet Res 2013;15(1):e7) doi:10.2196/jmir.2094

KEYWORDS

Information storage and retrieval; Medical Informatics; Education, professional

Introduction

Web-based learning repositories allow mass storage, management, and search and retrieval of data for both staff and students. Material can be stored in a variety of formats and easily shared across diverse user groups. This is particularly valuable in health profession programs in which the requirement for currency demands that resources are available for regular academic review and are visible to students and their clinical educators [1,2]. Spacious electronic repositories enable a comprehensive body of resources to be accessed easily by geographically dispersed users. Repository designers can also individualize access for critical review of resources and enable ongoing resource updating and refinement [3,4]. Health profession students in workplace practice benefit from access to learning resources that promote learning experiences and maximize movement toward learning targets.

Against a tradition of hard copy learning resources [5], Internet-based learning resources are rapidly augmenting or replacing other forms of information storage and sharing [6]. Students today enjoy social immersion in Internet-based connections and embrace Internet access to learning material [7]. Students of health professions report valuing access to online resources [8,9], with students in many programs relying at least in part on Internet-based resources for learning support [2].

A 2008 systematic review examined the effects of Internet-based learning compared with either no intervention or non-Internet-based learning, (eg, classroom instruction) for students of health professions. The review identified that Internet-based learning and traditional teaching methods appeared to have similar effects with regard to student satisfaction, knowledge, behavior, and patient outcomes [6].

Little is known about the value of Internet-based resources for health profession students in supporting workplace practice: which resources they find valuable, how and when they prefer to access them, and the form of resources they prefer to access. As educators determine the resources made available to students, they may benefit from learning more about the needs and preferences of users. Packaging learning material to maximize its appeal and availability to learners can potentially increase engagement and uptake.

This study was designed to examine (1) when and why health professional learners access resources held in an online repository, (2) which resources they preferentially access, (3) what resources they find most helpful in supporting their clinical placements and development of practice competencies, and (4) whether uptake is changing across time.

Methods

The Physeek digital repository is an online, keyword-searchable, repository of learning resources (see Multimedia Appendix for Screenshots). It was developed originally for students undertaking the 4-year undergraduate physiotherapy degree at Monash University, Melbourne, Australia. It enables remote access to learning resources to students in workplace practice. Students undertake 39 weeks of clinical learning during the third and fourth years of their studies. Physeek provides an appropriate model for Web-based learning repositories because it allows academic staff to create, store, and manage educational content in electronic format. It can be used by anyone with permission and Internet access. Third- and fourth-year students can search Physeek by using a number of search strategies (ie, subject category, keyword, year level, resource author, and resource type), whereas first- and second-year students who are not currently undertaking clinical placement can access Physeek resources through a generic learning management system (LMS), specifically Blackboard. The resources are intended to support and advance practice competency. They include lecture notes, practical demonstration videos, self-directed learning modules, and practical class pre-readings. Access to the repository was provided to clinical educators affiliated with the Monash program so that they were aware of course content and the expected level of student knowledge and skills. Educators also appreciated the opportunity to compare their own knowledge and beliefs to the current concepts taught to students.

The study was approved by Monash University Human Research Ethics Committee, approval number CF10/3439-2010001817.

Physeek Usage Audit

Usage of the Web-based repository Physeek was audited through 2 corresponding academic semesters in 2009 and 2010. Reports for students at each year level (1-4) of the bachelor of physiotherapy degree were generated. Reports generated were (1) number of online resources accessed per day in 2010 (years 3-4); (2) number of each type of resource accessed (eg, lecture slides, practical videos, unit guides, self-directed modules, practical pre-readings [years 3-4]); (3) number of resources accessed during business hours (9 am to 5 pm) and number accessed outside business hours (years 1-4); (4) session length of each Physeek log-on (years 1-4); (5) video quality (bit rate) of each video accessed (all year levels grouped). Each video resource was loaded onto the repository in 3 different quality levels based on bit rate measured in kilobit per second (kbit/s): low quality (56 kbit/s or 150 kbit/s depending on program used), medium quality (256 kbit/s or 400 kbit/s), and high quality (512 kbit/s or 720 kbit/s). The highest bit rates allowed the greatest image resolution. Videos were uploaded onto the repository at different qualities to allow for differences in bandwidths available to students. Download time is affected by file size, and we were curious to know what file sizes students preferentially selected to download (faster downloads with poorer image resolution or slower downloads with higher resolution).

Data Analysis

The number of online resources accessed per calendar day in 2010 were plotted against the third- and fourth-year clinical timetable.

The number of each type of resource accessed (lecture slides, practical videos, unit guides, self-directed learning modules, and practical pre-readings) in 2009 were compared to the number accessed in 2010. Chi-square tests were used to investigate if there were significant differences in uptake between 2009 and 2010 for each type of resource.

The number of resources accessed during business hours and outside business hours during 2010 were expressed as a percentage of total resource accessed for each year level (1-4). Data from first- and second-year students were included in this analysis to investigate if there were differences in the way that clinical education-based and campus-based students accessed resources. A Chi-square test was used to determine if usage within and outside business hours was significantly different.

The sum of video downloads at each quality level (low = 56 kbit/s or 150 kbit/s; medium = 256 kbit/s or 400 kbit/s; high = 512 kbit/s or 720 kbit/s) was expressed as a proportion of the total number of video downloads with associated 95% confidence interval (95% CI).

Descriptive statistics, including means, standard deviations (SD), and range were calculated to describe session length of Physeek log-ons by year level. A Kruskal-Wallis test was used to investigate differences in mean session length between each year level.

Questionnaire

Participant Inclusion/Exclusion Criteria

Students enrolled in fourth year of the bachelor of physiotherapy degree (ie, those with previous clinical experience) in the year 2011 were invited to participate in the online questionnaire. Other inclusion criteria were that they were adults (>18 years), fluent in English, and had completed clinical placements in 2010. Third-year students were not included because they had not previously undertaken clinical placement and did not have experience of learning and practice in the workplace.

Recruitment

A bulk email invitation was sent to all fourth-year students. To reduce student perceptions of coercion, the email was sent by a research assistant who was independent of the course. The email included an explanatory statement and a hyperlink to an anonymous online questionnaire.

Data Collection and Analysis

The questionnaire consisted of 6 statements about Physeek usage and its impact on learning. Responses to the first 5 statements (eg, “I have found Physeek helpful in revising practical skills for clinical placement”) were graded on a 5-point Likert scale. Responses to items were assembled using descriptive statistics.

Focus Group

Participant Inclusion/Exclusion Criteria

Eligible participants included fourth-year physiotherapy students (with previous clinical experience), aged over 18 years, and enrolled for study in 2011. Participants were excluded if they did not participate in clinical placement during 2010.

Recruitment

A convenience sample was employed to recruit eligible participants. Fourth-year students were emailed an invitation to participate by an independent research assistant. The first 18 students to reply were included in focus groups. Participants were randomly allocated to 3 groups of 6 to maximize the opportunity for individual participation [10].

Data Collection

Focus group discussions were facilitated by an independent researcher who was not involved with teaching in the bachelor of physiotherapy program. Group discussions were prompted by using a list of questions developed based on a review of the literature and designed to facilitate participant interaction. Each focus group was of approximately 30 minutes duration.

Questions were intended to generate general opinions about the delivery of learning resources that support workplace practice and, in particular, the usefulness of the Physeek database as a vehicle for accessing learning resources. Sessions were audiorecorded and transcribed verbatim before coding and thematic analysis.

Data Analysis

Transcripts for each focus group were independently coded by 2 researchers and themes were independently devised by using the principles of thematic analysis [10]. Thematic analysis indicated saturation of ideas after 3 focus groups and further participant recruitment was halted. To ensure accuracy of transcripts, a third independent researcher compared transcripts to the original audiorecording. Two researchers then compared and discussed the coded and themed transcripts for similarities and differences. This process was repeated for the transcripts of each focus group. When there was disagreement concerning major themes after 2 rounds of comparative analysis and discussion, a third independent researcher was consulted. After agreement on key themes was reached, 2 authors independently reviewed transcripts for suitable validating quotes. Quotes were also extracted in which they corroborated results of the usage audit and online questionnaire.

Results

Participants included in the usage data were third-year students (n=62) and fourth-year students (n=57) during 2009, and third-year students (n=48) and fourth-year students (n=64) during 2010. The total number of participants over both years was 231.

Physeek Usage Audit

Daily Frequency of Physeek Usage During 2010

Physeek usage by day from January 1 to September 29, 2010, is shown in Figure 1 (third-year students) and from February 1 to May 31, 2010, in Figure 2 (fourth-year students). These periods were chosen to be displayed because they represent a period of relatively intense study load and the time of highest Physeek resource uptake for each year level. Physeek usage for third-year students peaked at 385 accesses per day on May 10, 2010, at the start of the examination period, and at 183 accesses on February 8, 2010, for fourth-year students (183 accesses) immediately before the campus-based preclinical week.

Figure 1. Daily Physeek usage by third-year students mapped to the semester timetable from April 1 to June 30, 2010.

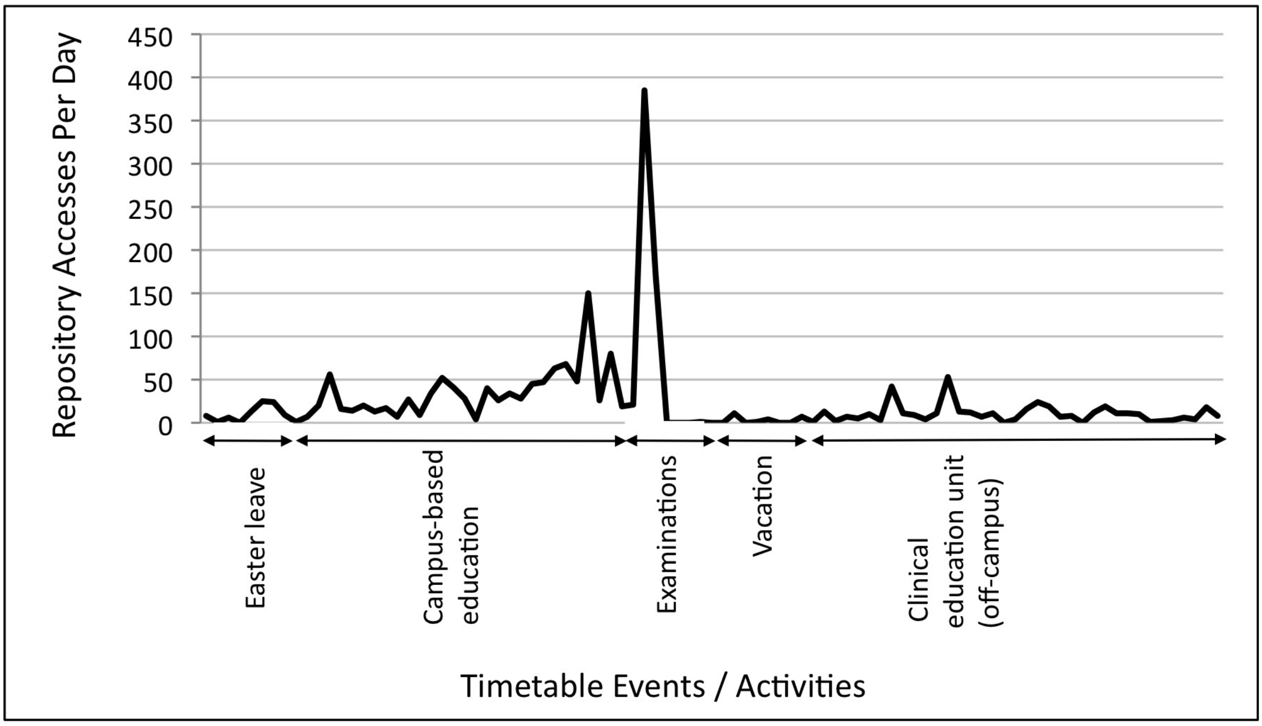
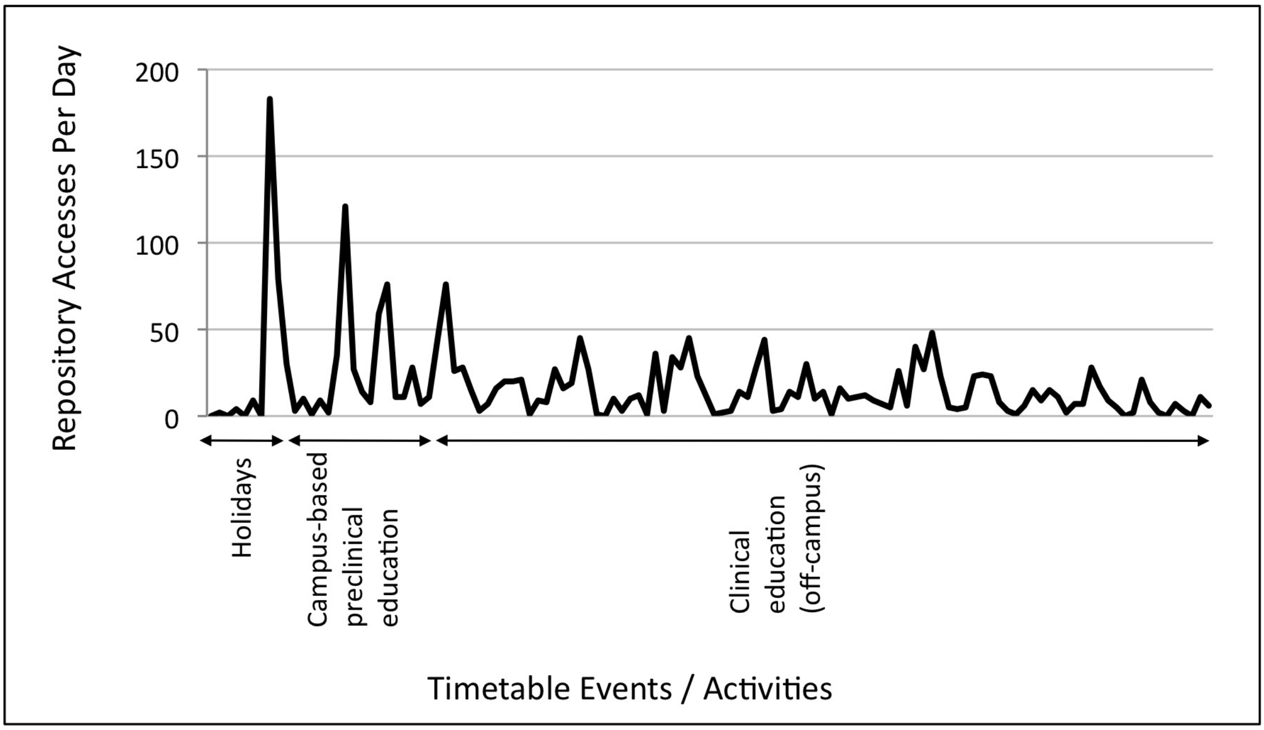


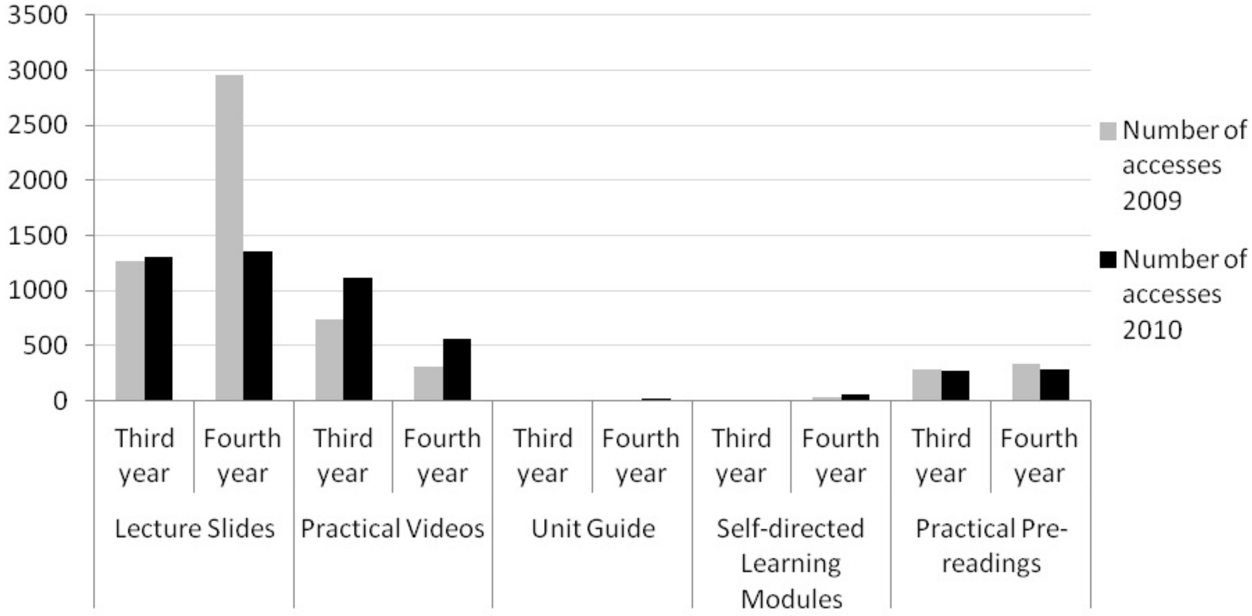
Figure 2. Daily Physeek usage by fourth-year students mapped to the semester timetable from February 1 to May 31, 2010.



Types of Resources Accessed From 2009 to 2010

Third- and fourth-year students both accessed lecture slides and practical videos considerably more than the other resources available on Physeek. There was minimal access of unit guides and self-directed learning modules, whereas a small number of practical pre-readings were accessed (Figure 3). When use was pooled across year levels, a Chi-square test showed a significant difference in the number of lecture slides accessed in 2009 compared to 2010 (*P*=.01), with a greater number being accessed in 2009. Conversely, there were significantly more practical videos accessed in 2010 compared to 2009 (*P*=.02). There were no significant differences in uptake of unit guides (*P*=.26), self-directed learning modules (*P*=.76), and practical pre-readings (*P*=.41) between 2009 and 2010.

Figure 3. Types of online resources accessed by third-year (n=62 for 2009, n=48 for 2010) and fourth-year (n=57 for 2009, n=64 for 2010) students.



Access During Business Hours Versus Outside Business Hours by Year Level 2010

All year levels (1-4) showed a preference for accessing Physeek during business hours (9 am to 5 pm), although the proportion of access outside business hours increased with year level. First-year students accessed Physeek during business hours 68% of the time (1315 separate accesses), second-year students 59% of the time (701 accesses), third-year students 55% of the time (1815 accesses), and fourth-year students 50% of the time (1639 accesses). There was a significant (*P*=.01) difference between year levels in the amount of resources accessed during business hours and those accessed outside business hours. A post hoc analysis using Pearson correlation coefficient was performed to examine the correlation between increasing year level (1-4) and proportion of time spent using Physeek outside business hours. This returned a significant Pearson correlation coefficient of 0.98 (*P*=.02).

Session Length by Year Level 2010

Average session lengths for each year level ranged from 46 minutes to 59 minutes, with average session times increasing with year level (Table 1). A Kruskal-Wallis test showed that there was a significant (*P*=.01) difference in mean session length between year levels. A post hoc analysis using Pearson correlation coefficient was performed to examine the correlation between increasing year level (1-4) and session length. This returned a significant Pearson correlation coefficient of 0.99 (*P*=.01).

Table 1. Physeek database session length (minutes) by year level for students during 2010.

Session information	First year	Second year	Third year	Fourth year
Mean (minutes)	46.9	52.2	55.2	59.6
Number of sessions	500	284	1295	1396
Standard error of the mean	1.1	1.5	2	2.8
Median	38.6	42	35.7	38.5

Video Bit Rate 2010

An audit of video bit rate preference demonstrated that all available bit rates were utilized by all year levels (1-4) when accessing practical videos. Across all year levels (1-4), there were 3590 video accesses. Low quality (56 kbit/s or 150 kbit/s depending on the program) video was accessed 16% (558/3560; 95% CI 13-19) of the time, medium quality (256 kbit/s or 400 kbit/s) was accessed 45% (1614/3560; 95% CI 43-47) of the time, and high quality video (512 kbit/s or 700 kbit/s) was accessed 39% (1388/3560; 95% CI 36-41) of the time.

Online Questionnaire

Thirty-nine students completed the online survey (81% response rate). A summary of student responses is presented in Table 2.

Table 2. Likert scale responses to survey regarding Physeek utilization.

Question	Response, n (%)[a]				
	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
I have found Physeek useful in revising practical skills for clinical placement	1 (2.6)	0 (0)	0 (0)	18 (46.2)	*20 (51.3)*
Physeek has increased the efficiency of searching for and locating clinical revision resources	1 (2.6)	0 (0)	0 (0)	10 (25.6)	*28 (71.8)*
Physeek has increased the time I have spent revising clinical skills	0 (0)	2 (5.1)	12 (30.8)	*15 (38.5)*	10 (25.6)
Physeek has increased the likelihood that I would seek a resource before its intended delivery	0 (0)	4 (10.5)	5 (13.2)	*18 (47.4)*	11 (28.9)
Access to resources via Physeek has enhanced my clinical performance	1 (2.6)	0 (0)	6 (15.4)	*21 (53.8)*	11 (28.2)
I have chosen not to attend a lecture or practical session in the past 12 months due to a knowledge that resources, ie, lecture notes, would be available on Physeek[b]	*20 (51.3)*	13 (33.3)	4 (10.3)	2 (5.1)	0 (0)

[a] Mode response for each question is represented in italics.

[b] Responses for this question correspond to never, rarely, sometimes, frequently, and always instead of strongly disagree, disagree, neutral, agree, and strongly agree, respectively.

Focus Groups

There was a general consensus among participants that the Physeek database provided an easier and more efficient means of accessing resources than traditional methods of delivery. Most students agreed that this provides an advantage in supporting clinical performance. However, student experiences varied with regard to the ways they utilized the repository, For example, a number of students viewed Physeek primarily as a “quick access” revision tool, but many also saw the potential for an online repository to be the primary mode of delivery for new or additional material. Similarly, students were divided on the future incorporation of technology and electronic education into health care practice. Results are presented for 3 key themes: (1) an online searchable repository is the preferred method of accessing learning resources efficiently in workplace practice; (2) make it bigger, faster, and easier; and (3) online repositories may be an effective tool to support lifelong learning and health care delivery. Supportive quotes demonstrating each of these themes are provided.

An Online Repository Is the Preferred Method for Efficient Access to Learning Resources

A digital repository such as Physeek cannot have a positive effect on learning and physiotherapy practice if students do not access its resources. Students varied somewhat in the extent to which they felt Physeek was useful to their clinical performance, but all agreed that Physeek was the easiest way to access resources needed for clinical revision. This comment from one student encapsulated the feelings of most participants when comparing Physeek to other means of accessing resources to support clinical education:

> *...you can’t even compare them...there’s just that instant access point [to the learning resources via Physeek] when you’re on clinic.*

There were a number of reasons why students preferred an online repository as opposed to other methods. For example, one topic that presented several times was the idea of efficient access to relevant resources:

> *Physeek’s been really useful because you don’t have to go and sift through...*
>
> *If...I didn’t know where to look for a piece of information, I could just search it and it would come up with a list of things that could possibly be useful.*

Not only did students feel that Physeek was an efficient way of accessing resources, they also believed that the resources could be accessed at a time that was clinically relevant:

> *...especially on clinics when I was doing outpatients, we’d know our patient the day before so then I’d go home that night and look at all the tests and everything that I could do, and think about all the treatment options...*

One of the other advantages of having an online repository raised by participants was linked to the idea of resource reliability and academic review. Many students felt that because resources were produced and updated by an academic organization (ie, the university, evidence-based resources), they were more confident in using them:

> *You know it’s reliable instead of like Google, at least stuff on Physeek is actually correct.*

Students Want a Repository That is Larger, Faster, and Easier to Use

The focus groups revealed a number of student recommendations for how an online repository might provide resources more effectively to facilitate clinical performance. These recommendations primarily focused on the idea that Physeek could be expanded, while improving the search functionality of the repository. Some students remarked that, at present, they had some difficulty in finding relevant resources using the Physeek search function:

> *...there are other times when I've just searched for something and the wrong information comes up.*

Several students suggested that mapping or browsing functions may make searching the database more accurate, while increasing the number of resources relevant returned for consideration:

> *If you typed "neuro" [neurological] it would...map it to everything that's related to neuro so say I had a neuro placement coming up, you could go, I don't know what I want to learn, but I'll just...search and see what resources they've got.*

Students also felt that a browse function may enhance their ability to search through resources, enabling them to look for cues regarding knowledge and skills that may be useful:

> *...you get a browse function so...you could just go look at musc [musculoskeletal] and just scroll through every musc thing we've done.*

Lack of technological prowess did not seem to affect opinions about Physeek. One of the key factors in facilitating uptake of Physeek resources may be related to limiting the technological knowledge and skills needed to use the database:

> *No, I hate technology and that's why I like Physeek, because it's just like Google. If I don't know what I'm looking for specifically, I can just type in MS [multiple sclerosis] and everything will come up.*

Online Repositories May Be an Effective Tool to Support Lifelong Learning and Health Care Delivery

There are potential advantages to integrating technology into health care practice. Students' attitudes toward lifelong learning and the capacity for an online repository to enhance this demonstrated the capacity for an online repository to improve individual health care delivery:

> *It would help if we still had access to it after we finish...because as a new grad, you're still going to be wanting to revise.*

These attitudes reflect a desire by the students to continue learning and advancing their profession after graduation:

> *If you don't embrace it, physio's just going to be left behind and everyone's just going to be doing what we did twenty, thirty years ago, nothing's going to change.*

Discussion

This paper presents the results of a multifaceted study examining student usage and perceptions of a Web-based digital repository designed to improve access to learning resources for health professional learners. This is the first study to examine students' usage of an online learning repository as a support for workplace practice and relate this to student experiences and preferences for learning resources. The findings of this study provide insight into how online repositories could be designed or utilized by educational institutions to enable maximal resource access and uptake for health professional learners.

Students saw the Web-based repository as the most efficient and the preferred source of learning resources, but also felt that Physeek could be further improved to make resource access even more efficient. This emphasis on efficiency is reinforced by comments that a major advantage of a Web-based repository is a reduction in the need to wade through irrelevant search yields. This may account for the rapid uptake of Physeek resources. When one finds useful information in a time-efficient way, this is likely to provide positive reinforcement to the exercise of looking for relevant material. This feedback loop encourages students to become active learners.

The results of the online questionnaire and usage audit suggest that Physeek may have had a positive impact on student study habits. Students reported that Physeek had increased the time they spent on clinical revision, and the likelihood that they would seek out resources before their intended use, again indicating that a Web-based repository may have made students more willing to engage in independent learning.

An audit of Physeek usage found that students were most likely to access online learning resources to satisfy immediate academic requirements, such as examination preparation or for gathering information needed on clinical placements. Physeek usage for both third- and fourth-year students peaked at times of increased study load (eg, third-year examinations and the fourth-year preclinical preparation period). These results are encouraging because they demonstrate a willingness by students to revise learned knowledge by using Web-based resources.

With the immediate access to knowledge and skills that an online repository provides, there is the potential for students to feel that they can rely on the repository as a source of information at the point of service delivery. All year levels in 2010 accessed resources through Physeek during business hours (9 am to 5 pm) significantly more than outside business hours. However, the first 2 years of study are primarily campus-based, allowing considerable time for computer access on campus during business hours, whereas both third- and fourth-year students accessed resources outside business hours more than 45% of the time while on placement. Revision session length for third- and fourth-year students was also typically greater than 55 minutes when logged on to Physeek.

This represents a considerable amount of time spent revising material on Physeek outside of clinical placement hours. Key student motivators for after-hours access are not clear, and may have been influenced by a lack of time during clinical hours, reduced Internet access, or because the use of mobile devices at the point of care is currently discouraged in most clinical environments. Another study limitation is that although there is no motivation for accessing the repository without utilizing its resources, behavior of this kind would impact the accuracy of the usage data.

It was clear that the availability of resources through a Web-based repository did not appear to have a negative impact on attendance at scheduled learning sessions. More than half (51%) of the students in the study reported they had never skipped a lecture or practical session in the past 12 months because the learning resources were also available on Physeek. Another 33% said that they rarely did so. Only 2 students (5%) said that they missed sessions frequently because the resources were available on Physeek. Students reported that they were more likely to use Physeek to enhance or broaden their knowledge rather than as an alternative to lectures or practical classes. This supports the findings of an observational study by Grabe and Christopherson [2], who found that providing supplementary online resources, in this case lecture notes, could be related to improved class attendance. In this study, an online repository appears to enable meaningful revision and consolidation of knowledge, as well as immediacy of information at the point of application. A repository such as Physeek may, therefore, enhance clinical learning and practice, as well as patient care.

Students overwhelmingly preferred lecture notes and practical videos as the learning resources they accessed online. This is in agreement with the focus group discussions in which students were almost unanimous in reporting that these were the resources most helpful to them. It may be students view these as critical information for professional performance, whereas self-directed learning modules and practical notes are outside the core curriculum. However, one group of students appeared not to realize that these non-core resources were available on Physeek. Because Physeek provides only a search function and not the capacity to browse topics, it is possible that students were not aware of all the resources available to them via the online repository. As suggested by a number of students in the focus groups, a browsing or cataloging system may overcome this limitation and increase uptake of different types of resources.

Students tended to prefer better quality video for practical skill revision, with the majority of videos accessed of medium or high quality. However, approximately one-fifth of video accesses in 2010 (n=588) utilized low quality video. It was suggested from the focus group transcripts that this may be a matter of convenience. For example, students who needed to access video quickly or at slow connection speeds might use lower bit rate video at one time, and preferentially utilize higher quality video when they had access to greater Internet bandwidth. A qualitative study by Blake [9] also found that it is this type of flexibility that attracts students to the use of online material. This supports one of the major themes from the focus group, of making the repository “larger, faster, and easier to use.” Students generally wanted more resources available through Physeek, while at the same time maintaining or improving the ease and efficiency of finding these resources.

As yet, no major modifications have been made to the repository; therefore, it appears that the uptake of Physeek resource access was primarily driven by growing awareness and familiarity with the repository, and an increasing student desire to access resources online. The results of Chi-square tests comparing uptake of each resource between 2009 and 2010 indicated that usage remained consistent over this time. Given that the technology for Physeek was implemented at the start of 2009 and students were still familiarizing themselves with using a Web-based repository, the significant amount of time spent on Physeek over the 2-year period (Table 1) is encouraging. These data will need to be re-evaluated as modifications are made to the repository based on student feedback. Further research could also be conducted to identify which aspects of the Physeek repositories’ properties and functionality has influenced its successful implementation, such as the particular tagging schema utilized allowing resources to be searched by category as well as per individual item, the consistent format of video resources allowing playback by the built-in media player, or the ability for other learning systems to link to resources within the Physeek repository.

Rather than providing only supplementary material, it is possible that Web-based repositories can be the primary means of accessing resources for health professional learners, as well as a tool for educating patients and accessing “practical” information. For example, Physeek may be useful in providing patient education materials that can be readily accessed at the bedside. When asked about possible further applications of the online repository, students pointed to the capacity to improve clinician recall and patient education by using interactive media, while also enhancing lifelong learning for health care professionals.

The results of this study indicate that today’s health professional students welcome the benefits of online learning resources because of their convenience and usability. This represents a transition away from traditional learning styles toward technological learning support, and may indicate a growing link between social immersion in Internet-based connections and learning styles. The true potential for Web-based resources to support student learning is as yet unknown. Many students expressed a desire to continue using this type of resource throughout their professional lives. With careful design, a Web-based learning repository is potentially a useful model to provide support for health professional students and graduates in workplace practice.

Acknowledgments

This research was funded in full by the Monash University Faculty of Medicine, Nursing and Health Science. This project has been approved by the Monash University Human Research Ethics Committee (MUHREC), project no: CF10/3439-2010001817.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Screenshot 1 - Example of the repository search yield display.

[JPG File, 61KB - jmir_v15i1e7_app1.JPG]

Multimedia Appendix 2

Screenshot 2 - Example of a repository item description.

[JPG File, 55KB - jmir_v15i1e7_app2.JPG]

References

1. Billings ME, Engelberg R, Curtis JR, Block S, Sullivan AM. Determinants of medical students' perceived preparation to perform end-of-life care, quality of end-of-life care education, and attitudes toward end-of-life care. J Palliat Med 2010 Mar;13(3):319-326 [FREE Full text] [doi: 10.1089/jpm.2009.0293] [Medline: 20178433]
2. Grabe M, Christopherson K. Optional student use of online lecture resources: resource preferences, performance and lecture attendance. Journal of Computer Assisted Learning 2008;24(1):1-10.
3. Brickell J, Kanuth M, Freeman V, Latshaw S, Larson C. Learning objects: resources for instruction. Clin Lab Sci 2006;19(3):184-187. [Medline: 16910236]
4. Ruiz JG, Mintzer MJ, Issenberg SB. Learning objects in medical education. Med Teach 2006 Nov;28(7):599-605. [doi: 10.1080/01421590601039893] [Medline: 17594550]
5. Premkumar K, Baumber JS. A learning resources centre: its utilization by medical students. Med Educ 1996 Nov;30(6):405-411. [Medline: 9217901]
6. Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ, Montori VM. Internet-based learning in the health professions: a meta-analysis. JAMA 2008 Sep 10;300(10):1181-1196. [doi: 10.1001/jama.300.10.1181] [Medline: 18780847]
7. Veale P. Prospective comparison of student-generated learning issues and resources accessed in a problem-based learning course. Med Teach 2007 May;29(4):377-382. [doi: 10.1080/01421590701509712] [Medline: 17786754]
8. Schimming LM. Measuring medical student preference: a comparison of classroom versus online instruction for teaching PubMed. J Med Libr Assoc 2008 Jul;96(3):217-222 [FREE Full text] [doi: 10.3163/1536-5050.96.3.007] [Medline: 18654658]
9. Blake H. Computer-based learning objects in healthcare: the student experience. Int J Nurs Educ Scholarsh 2010;7:Article 16. [doi: 10.2202/1548-923X.1939] [Medline: 20597855]
10. Rice P, Ezzy D. Qualitative Research Methods. South Melbourne: Oxford University Press; 2009.

Abbreviations

kbit/s: kilobit per second
LMS: learning management system

Edited by G Eysenbach; submitted 07.03.12; peer-reviewed by R Ajjawi; comments to author 29.04.12; revised version received 05.05.12; accepted 05.10.12; published 16.01.13.

Please cite as:
Maloney S, Chamberlain M, Morrison S, Kotsanas G, Keating JL, Ilic D
Health Professional Learner Attitudes and Use of Digital Learning Resources
J Med Internet Res 2013;15(1):e7
URL: http://www.jmir.org/2013/1/e7/
doi:10.2196/jmir.2094
PMID:23324800

Original Paper

Online Information Exchanges for Parents of Children With a Rare Health Condition: Key Findings From an Online Support Community

Florin Oprescu[1], MD, MPH, MBA, PhD; Shelly Campo[2], PhD; John Lowe[1], MPH, DrPH; Julie Andsager[2], PhD; Jose A Morcuende[2], MD, PhD

[1]University of the Sunshine Coast, Faculty of Science, Health, Education and Engineering, School of Health and Sport Sciences, Maroochydore, Australia
[2]University of Iowa, Iowa City, IA, United States

Corresponding Author:
Florin Oprescu, MD, MPH, MBA, PhD
University of the Sunshine Coast
Faculty of Science, Health, Education and Engineering
School of Health and Sport Sciences
Locked Bag 4
Maroochydore, 4558
Australia
Phone: 61 754594639
Fax: 61 754594880
Email: foprescu@usc.edu.au

Abstract

Background: The Internet provides new opportunities for parents of children with difficult illnesses and disabilities to find information and support. The Internet is particularly important for caregivers of children with special needs due to numerous health-related decisions they face. For at-risk populations, online support communities can become key settings and channels for health promotion and communication.

Objective: This study is an initial exploration of the information-seeking and information-provision processes present in an online support community, which is an area of opportunity and interest for Internet-based medical research and practice. The aim of this study was to explore and describe information-related processes of uncertainty management in relationship to clubfoot. Specifically, the study explored interpersonal communication (information seeking and provision) in an online support community serving the needs of parents of children with clubfoot.

Methods: The study population consisted of messages posted to an online community by caregivers (parents) of children with clubfoot. The theoretical framework informing the study was the Uncertainty Management Theory (UMT). The study used content analysis to explore and categorize the content of 775 messages.

Results: Women authored 664 of 775 messages (86%) and men authored 47 messages (6%). Caregivers managed uncertainty through information seeking and provision behaviors that were dynamic and multilayered. The ratio of information-seeking messages to information-provision responses was 1 to 4. All five types of information-seeking behaviors proposed by Brashers' schema were identified, most of them being correlated. Information seeking using direct questions was found to be positively correlated to self-disclosure (r=.538), offering of a candidate answer (r=.318), and passive information seeking (r=.253). Self-disclosure was found to be positively correlated to provision of a candidate answer (r=.324), second-guessing (r=.149), and passive information seeking (r=.366). Provision of a candidate answer was found to be positively correlated with second-guessing (r=.193) and passive information seeking (r=.223). Second-guessing was found to be positively correlated to passive information seeking (r=.311). All correlations reported above were statistically significant (P<0.01). Of the 775 messages analyzed, 255 (33%) identified a medical professional or institution by name. Detailed medical information was provided in 101 (13%) messages, with the main source of information identified being personal experience rather than medical sources.

Conclusion: Online communities can be an effective channel for caregivers, especially women, to seek and offer information required for managing clubfoot-related uncertainty. To enhance communication with parents, health care institutions may need to invest additional resources in user-friendly online information sources and online interactions with caregivers of children with special illnesses such as clubfoot. Furthermore, explorations of information-seeking and information-provision behaviors in online communities can provide valuable data for interdisciplinary health research and practice.

(J Med Internet Res 2013;15(1):e16) doi:10.2196/jmir.2423

KEYWORDS

Online social support; Online support communities; Clubfoot; Uncertainty management; Health communication; Health information seeking and provision

Introduction

The Internet is particularly important for caregivers of children with special needs due to numerous health-related decisions they face [1]. For at-risk populations, online support communities can become key settings and channels for health promotion and communication [2-4]. However, limited data are available on the uncertainty management behaviors exhibited in online support communities for caregivers of children with clubfoot [5]. In this context, it is important to explore and analyze information seeking and information provision as uncertainty management behaviors [6].

A parent or caregiver experiences uncertainty when caring for a child affected by illness or disability [7-9]. Clubfoot is a developmental disability affecting the lower limb with an incidence of approximately 1 per 1000 live births [10]. Given its visual nature, with one or both feet turned inwards at birth, clubfoot can be a major source of uncertainty for the parents of the children affected after diagnosis and during treatment [5,11]. Uncertainty can also manifest because of the relapses that may occur after treatment [12-14]. Furthermore, available clubfoot treatment options vary and their efficacy depends on the experience of the health care provider [15-17]. Due to the limited time available for consultation with health care professionals and the relative rarity of the condition, caregivers may be likely to turn to online support communities to manage their illness-related uncertainty.

Uncertainty is defined as a psychological state characterized by insecurity and lack of clear information [18]. Similar to other serious health conditions, the parents of children with clubfoot face a complex situation: information about treatment may be unavailable, information about treatment may be inconsistent, and there may be insecurity related to limited general or specialized knowledge about what causes clubfoot and what the long-term effects of the condition are [19,20]. Further sources of uncertainty may include potential stigma of the health condition, child development (including long-term effects of illness), as well as selection and effectiveness of treatment options [21,22]. Faced with such a condition, the parent/caregiver may employ various behaviors to manage uncertainty. However, there are gaps in knowledge regarding uncertainty management behaviors of caregivers, especially in online environments.

Uncertainty Management Theory

The Uncertainty Management Theory (UMT) addresses a number of ways in which impacted individuals attempt to reduce, maintain, or increase their level of uncertainty [18]. Individuals can manage uncertainty (theirs and others') through appraisal processes and behavioral responses. Situation appraisals can influence the magnitude of uncertainty and its impact on the individual [23]. Behavioral responses may include information seeking and information provision, the focus of this study. In addition to fulfilling the need for acquiring and sharing information, exchanging information online may help parents feel more secure in their role as a caregiver of a child with clubfoot and validated as members of the community [22,24,25].

Information Seeking and Uncertainty Management

Information is one of the tools available to manipulate uncertainty [26]. Information seeking may include question asking, self-disclosure, offering a candidate answer, second-guessing, and passive information seeking [18]. However, the distribution of information-seeking behaviors and their correlations is yet to be documented.

In health-related situations, the high stakes may lead to a large number of venues and styles of information seeking to ensure that the individual gains access to as much information as possible regarding a certain health issue [27,28]. The Internet is a venue of increasing importance for health information exchanges. A report from the Pew Internet Research Institute [29] indicates that there has been an explosion of health-related information online, both in terms of production and consumption, because of the increase in Internet access (74% of American adults use the Internet). The Pew survey found that 80% of American Internet users have searched online for health information and that 70% of American Internet users who are also caregivers for an ill person look for health information online. Of those surveyed, 40% of caregivers indicated that online health information had been helpful [29].

Parents on the Internet

Most clubfoot cases are diagnosed at birth, and some of the new parents may not be able to receive and process sufficient relevant information during postnatal care [30]. As a result, new parents may use the Internet as their major source of health information [31]. It has been documented that most parental Internet users are women [32,33], their mean age is less than 35 years old, and many are first-time parents [34]. However, their patterns of information seeking and provision, especially in online communities are unknown.

Given the increased availability of medical information, more parents want to be part of treatment-related decision making. The increased availability of medical information can enable them to become informed medical consumers [35]. Parents may become dissatisfied with the medical encounter if not given sufficient attention or access to informational resources [36] and leave the consult appointment feeling that they can search for, find, and use medical information that is superior to the recommendations of the physician [35]. To manage uncertainty regarding the health care provider, parents may attempt to identify and recommend the physicians and medical institutions that provide good care while steering away from those who provide less than satisfactory care [37]. Online messages that include identifying information of health care professionals and institutions may be used to manage the uncertainty of the recipient (reduce, increase, or maintain) regarding selection of the medical care provider [38]. Research exploring whether informational exchanges within online support groups mention health care providers and institutions is still in its infancy.

The use of the Internet as an information and support source is of particular importance for parents of children with illnesses [31]. A child affected by illness or disability increases the need for parental support and information [31], and parents who need to care for such a child often become active seekers of information [28]. These parents may attempt to find as much information as possible about the health condition of interest including the best treatments, doctors, and facilities available [37]. Information may be sourced using search engines, websites recommended by friends, advertisements in parental magazines, and online support groups [31,39]. Data describing types of medical information provided in online communication and their origin in the context of clubfoot or similar congenital disabilities are currently unavailable.

Information Seeking in Online Support Groups

Repeat visits to monitor online groups and obtain updates may be considered as part of the ongoing process of managing uncertainty by having the most up-to-date information and confirming that the information acquired so far is still valid [40]. Online support groups have a number of advantages such as 24/7 availability, lack of geographical barriers, a greater degree of anonymity, and ability for people to carefully read and compose messages [41-43]. Given the relative rarity of the clubfoot condition, another advantage is the ability to find or even meet other parents who face similar challenges in caring for their children with clubfoot. Some disadvantages may include lack of physical contact, potential of negative experiences, and lack of information quality control mechanisms [44,45]. The lack of quality control in online communities, especially when discussing medical issues, may result in information that could be conflicting, misleading, or even invalid [46,47]. Thus, it is important to identify the sources of medical information exchanged in a user-managed online community that is not monitored by health care professionals.

Study Aim

The aim of this study was to explore and describe information-related processes of uncertainty management in relationship to clubfoot. Specifically, the study explored interpersonal communication (information seeking and provision) in an online support community serving the needs of parents of children with clubfoot. The setting for the study was a user-managed online community for the parents of children with clubfoot. This study adds to the body of research seeking to understand better how parents use online support communities to seek information and to manage uncertainty when caring for children with rare health conditions.

Methods

Study Population

The study population was represented by messages posted to the oldest and largest Yahoo-based user-managed online support community dedicated to informational and social support needs of parents of children with clubfoot. In 10 years, the group members exchanged over 76,000 messages. The group had over 2300 members with approximately 20 new members joining each week. Members posted over 50 new messages every week. Active members posted messages to the community, either seeking information or providing information and other types of support.

Study Sample and Methodology

The study sample consisted of randomly selected messages. The sampling methodology was systematic random sampling [48]. The sampling rate was 100 (every 100th message was collected) starting with a randomly selected message posted in the early days of the group. Researchers collected, coded, and analyzed 775 messages posted between January 2000 and December 2008. The study methodology received ethical approval from the University of Iowa’s Institutional Review Board, and the group owner and administrator provided permission to conduct the research.

The study used content analysis to analyze and code messages. Content analysis is an accepted method to study exchanges in online support communities [42,49,50]. Descriptive variables included gender of author, type of message, and intended recipient of the message. These variables were recorded separately during the data collection and deidentification process. The gender of the author was recorded as female, male, and unknown. The type of message was recorded as original message if it was the first message in the thread or response to a previous message if not the first message in the thread. The intended recipient was recorded as individual if the message was clearly addressed to an individual (ie, addressed by name) or group if the message was not addressed to a specific person.

The types of information-seeking behavior were coded following the schema proposed by Brashers [5]: question-asking, self-disclosure, offering a candidate answer, second-guessing, and passive information seeking.

Messages with identifying information were those that included a clear identification of a facility or medical professional, either by name or by a description. These messages were categorized based on the expressed overall experience of the poster in: negative (“I have to admit I have been left to feel rather neglected through this recent experience with the hospital.”), neutral, or positive (“We could not be happier with our doc. He is outstanding and caring...”).

Messages including information about diagnoses, symptoms, regular, and alternative treatments, relapses, and other medical issues were coded as messages dealing with medical information. The source of medical information was classified as personal if it was based on the author's experience, medical professional if a health care provider was mentioned as source, hospital or institutional website if a link was provided, and medical textbook or journal if the title was provided.

Statistical analyses, including intercoder reliability, were conducted using the SPSS software package. To compute intercoder reliability, out of the analytical sample, 15% (N=116) of messages were randomly selected and independently coded by 2 coders following the methodology proposed by Neuendorf [48] for intercoder reliability computations. Both coders were trained using 30 messages that were not included in the analytical sample. Both coders maintained coding notes as sources of data for thematic analysis using a naturalistic inquiry approach based on the constant comparison method [51]. During peer debriefing, the coders discussed 14 messages where coding disagreements occurred, and the final coding included the consensus opinion. To ensure validity, disagreements were documented in the coding notes prior to achieving consensus in order to reduce potential biases [49]. Krippendorff's alpha coefficients were computed for each variable. Krippendorff's alpha ranged between .84 and .98. A Krippendorff alpha above .80 was considered acceptable [48,52].

Results

Out of 775 messages coded, women posted 664 (86%) and men posted 47 (6%) of the messages. The gender of the author was unknown for 64 (8%) of the messages. 620 out of 775 messages (80%) were replies to a previous message, and 155 (20%) were initial messages for a ratio of 4:1, indicating an average number of four responses for each original message posted to the board. The intended recipient was an individual in 559 (73%) of the messages, while in 210 (27%) of the messages the intended recipient was the online support group as a whole.

The most frequent information-seeking behaviors were direct questions in 196 (25.3%) messages, followed by self-disclosure in 116 (15%) messages. Other types of information-seeking behaviors were identified as offering a candidate answer in 33 (4%) messages, passive information seeking in 31 (4%) messages, and second-guessing in 3 (0.4%) messages. Five Pearson product-moment correlations were conducted to determine the correlation between various types of information-seeking behavior. Information seeking using direct questions was found to be positively correlated to self-disclosure, offering of a candidate answer, and passive information seeking. Self-disclosure was found to be positively correlated to provision of a candidate answer, second-guessing, and passive information seeking. Provision of a candidate answer was found to be positively correlated with second-guessing and passive information seeking. Second-guessing was found to be positively correlated to passive information seeking. All correlations reported were significant at the $P<.01$ level (2-tailed); see Table 1 for more detail.

Table 1. Correlations between types of information-seeking behaviors.

	Self-disclosure	Candidate answer	Second-guessing	Passive
Direct question	.538 [a]	.318 [a]	.059	.253 [a]
Self-disclosure		.324 [a]	.149 [a]	.366 [a]
Candidate answer			.193 [a]	.223 [a]
Second-guessing				.311 [a]

[a] Correlation was significant at the .01 level (2-tailed).

Of the 775 messages analyzed, 255 messages (33%) included names of health care professionals or institutions. 84 of the 255 messages (33%) provided detailed comments about the named health care professionals or institutions. Of the 84 detailed comments related to health care professionals and institutions, 54 (64%) were positive. Community members strongly urged the recipient to seek a second opinion or to change the health care provider in 30 of the 84 messages (36%).

Of the 775 messages analyzed, 101 messages (13%) included detailed medical information. Of the 101 messages providing medical information, 45 messages (45%) addressed bracing (special shoes the child wears after the casting is completed), 13 messages (13%) provided general information about clubfoot and the treatment options, 11 messages (11%) addressed relapses, and 11 messages (11%) addressed casting issues. The sources of information identified were personal experience in 60 messages (60%), followed by a medical professional in 20 messages (20%), and medical textbooks/journals or hospitals/institutional websites in 5 messages (5%).

The analysis of coding notes indicated that information-seeking messages were generative, dynamic, multilayered, and repetitive. Information-seeking messages were likely to generate information-provision messages. While an exact count was not possible, most information-seeking messages seem to originate from new mothers who recently joined the community following the clubfoot diagnosis or relapse. Information-seeking messages reflected fear and anxiety. They were dynamic in content with questions ranging from diagnosis, treatment options, treatment problems, bracing problems, and later on relapses. Finally, to elicit information from community members, those in information-seeking mode employed multiple layers of information-seeking behaviors.

A majority of information-provision messages seemed to be authored by parents during or after the treatment. These messages often included a combination of information and emotional support directed to the receiver. During the early years of the community, some members used the content of previous well-written messages to answer new inquiries. During the later years of the community, given the repetitive nature of inquiries, some of the more experienced community members created and maintained a list of “frequently asked questions” and a collection of pictures to help respond more effectively to information requests.

The willingness to identify, generate, and use information needed to manage uncertainty was observed in both information-seeking and information-provision messages. The authors of information-seeking messages seemed to have as a primary objective the management of their own uncertainty. The authors of information-provision messages seemed to have as their primary objective the management of the recipients’ uncertainty.

Discussion

The study explored interpersonal communication (information seeking and provision) in an online support community serving the needs of parents of children with clubfoot. The discussion sections include: (1) online group characteristics, (2) information-seeking behaviors, (3) messages identifying health care providers, and (4) messages providing medical information and the sources of medical information. Females posted a large proportion of the messages. The ratio of initial messages to responses was 1:4. All five types of information-seeking behaviors proposed by Brashers’ schema were identified, most of them being correlated. Two thirds of the messages mentioned a medical professional, with one in three of those messages providing detailed comments. One in seven messages included medical information, with the main source of information specified being personal experience.

Online Group Characteristics

During the study period, the messages posted addressed various issues related to caring for a child with clubfoot. The majority of active members were mothers of children with clubfoot indicating that they are the parent more likely to actively engage in online support groups to find and share information about clubfoot. This supports other literature that suggests mothers take the main responsibility for the health care of the family in general and of the children in particular [37,39,53]. It also suggests that women may be more likely than men to actively use support groups [54]. This finding may inform future health communication initiatives and ensure that messages directed to parents are tailored to the predominantly female audience.

Information-Seeking Behaviors

The most frequently used types of information-seeking behavior in the online community were direct questions and self-disclosure, both of them linked with attempts to manage uncertainty. Furthermore, community members tended to use combinations of information-seeking behaviors designed to generate responses from the community and to build trust. This may allow them to manage multiple layers of interconnected uncertainties: both knowledge-related and interpersonal-related [20]. Information-seeking behaviors were the starting point for exchanges of information in online communities, creating conditions for combining knowledge from multiple sources such as individuals, health care professionals, and Internet-based resources [55].

Messages Identifying Health Care Providers

According to UMT, individuals use information and other means to increase, maintain, or decrease their uncertainty levels [5]. The results suggested that the main purpose of the clubfoot online community is to allow caregivers to request and exchange information and other types of support. Much of the information provided seems to originate from the personal experience of caring for a child with clubfoot. Such information, in addition to providing solutions to various issues encountered over the course of the treatment, also included names of health care professionals and accounts of hospital visits. Two thirds of the comments regarding physicians and clubfoot treatment were positive. Message distribution suggests that the purpose of the community was not to serve as a scoreboard for physicians or medical care institutions. However, where necessary, group members will urge for second opinions or even for changing medical care providers if the accounts of medical visits do not fit standards of treatment or if some members of the message board had negative experiences with a particular physician or hospital.

Medical Information

The fact that hospital and institutional websites were rarely the source of medical information posted in online message boards is a key finding since health care information is one of the top subjects that people are looking for on the Internet, with over 12.5 million searches per day focusing on health issues [35]. Additionally, more and more parents want to have as much information as possible about the health care and condition of their child [27,28].

One major reason for illness-related uncertainty is the lack of clear, accurate, and complete information from a trusted source [20,22]. Hospitals and medical care providers represent a trusted source of medical information [35], and yet it is important to note the limited reliance on and reference to medical information available on the websites of medical institutions. This may have three potential explanations: the information is not available, reliable information is difficult to find, existing information is difficult to understand, or physicians do not encourage their patients to take advantage of the existing reliable Internet-based information. While some hospital websites may provide clear and accurate information, caregivers rarely mentioned this information source.

The quality and effectiveness of medical information on the Internet is an issue of concern for health communication scholars and health care practitioners [31,56]. For enhanced health communication, physicians may need to become familiar with high-quality and reliable information websites. This will allow them to increase the quality of information available to parents and to manage their uncertainty in the context of clubfoot care and other similar health conditions. Therefore, it is suggested that medical institutions dedicate additional resources needed to create high-quality online resources providing information on various aspects of the health condition [57]. The existing medical institutional websites in general are perceived as less attractive and difficult to use [58]. Furthermore, in the context of clubfoot, it is suggested that health professionals recommend reliable websites to their clients during the medical encounter as an Internet prescription [44]. This may increase client satisfaction and efficiently use the available consultation time by answering critical care questions and directing clients to online resources such as online support communities [59].

In this context, both medical encounters and online support communities could be efficient as diffusion mechanisms for medical information. The online environment appears to offer excellent opportunities for health care professionals and health communication professionals to provide high-quality medical information to caregivers who are in information-seeking mode [57]. The online environment also offers opportunities for innovative interdisciplinary research that can use information technology to bridge gaps between nursing, medicine, and health communication among others [43,45,57,60].

Study Limitations and Future Research

First, this study was limited to an online support community dedicated to caregivers of children with clubfoot. Thus, generalizations to other conditions may not be possible. Future studies may need to explore and compare information-exchange processes in online communities dedicated to other conditions. Current advances in online content research such as natural language processing may be appropriate for analysis and comparison of large datasets across health conditions and disciplines [60].

Second, it is important to note that the members of any online community differ in their levels of participation [61]. There are members who write messages frequently, members who write occasionally, and members who only read messages. The data collected in this study did not allow for such a categorization of participants. Future research may need to examine online information exchanges while considering participation frequency and length. Furthermore, we do not know information-seeking behaviors for those who did not join the group. It may be that some caregivers do not use this community for various reasons and their information-seeking behaviors may be different [62]. Physician-to-patient interactions are one-to-one interactions, while interactions in the context of online communities are one-to-many and many-to-one. The online environment may allow one-to-many and many-to-one interactions to be relatively easy, efficient, and effective by combining the knowledge and intelligence of the group members and preserving the output in a written and persistent format. This is an advantage of online communities that deserves more attention and that will require further interdisciplinary research, for example, modeling the communication pathways and individual involvement in the online exchanges in accordance with the research pioneered by Bambina [63]. Such research could provide direction for additional ways that health care professionals could become more involved in the fabric and activities of user-managed online communities dedicated to caregivers of children with special health conditions. Finally, future research could employ collaboration between members of online communities and scholars from various disciplines [64,65].

Conclusion

There is an increasing interest among health practitioners and scholars in the Internet-based behaviors of parents of ill children because of the need to increase effectiveness of health communication initiatives. This paper described information exchanges in the context of an online support community created and managed by parents of children with clubfoot. Caregivers seeking and providing information in online communities can fill a critical gap through communication processes that are timely and relevant. To enhance communication with parents, health care institutions may need to invest additional resources in user-friendly online information sources and online interactions with caregivers of children with special illnesses such as clubfoot. Furthermore, explorations of information seeking and provision behaviors in online communities can provide valuable data for interdisciplinary health research and practice.

Acknowledgments

The study was supported by a dissertation fellowship from the University of Iowa. The funder had no input in the design of the study, the review, or approval of the manuscript. Coding support was provided by Saloni Nayar, communication specialist with the National Institutes of Health. Analytical support was provided by the Statistics Outreach Center at the University of Iowa. Editorial support was provided by David Duncan from the Engage Research Lab at the University of the Sunshine Coast.

Conflicts of Interest

JM is Medical Director of Ponseti International Association, dedicated to promoting conservative treatment of clubfoot using the Ponseti method. FO worked as a graduate research assistant for the Ponseti International Association and is currently a volunteer member of the association.

References

1. Knapp C, Madden V, Wang H, Sloyer P, Shenkman E. Internet use and eHealth literacy of low-income parents whose children have special health care needs. J Med Internet Res 2011;13(3):e75 [FREE Full text] [doi: 10.2196/jmir.1697] [Medline: 21960017]
2. Gold J, Pedrana AE, Stoove MA, Chang S, Howard S, Asselin J, et al. Developing health promotion interventions on social networking sites: recommendations from The FaceSpace Project. J Med Internet Res 2012;14(1):e30 [FREE Full text] [doi: 10.2196/jmir.1875] [Medline: 22374589]

3. Mauri M, Cipresso P, Balgera A, Villamira M, Riva G. Why is Facebook so successful? Psychophysiological measures describe a core flow state while using Facebook. Cyberpsychol Behav Soc Netw 2011 Dec;14(12):723-731. [doi: 10.1089/cyber.2010.0377] [Medline: 21879884]
4. 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]
5. Paulsen-Miller M, Dolan LA, Stineman A, Morcuende JA. Understanding the educational needs for parents of children with clubfoot. Orthop Nurs 2011;30(4):273-8; quiz 279. [doi: 10.1097/NOR.0b013e3182247c15] [Medline: 21799385]
6. Cutrona CE, Suhr JA. Controllability of stressful events and satisfaction with spouse support behaviors. Comm Research 1992;19:154-174.
7. Mishel MH. Reconceptualization of the uncertainty in illness theory. Image J Nurs Sch 1990;22(4):256-262. [Medline: 2292449]
8. Pietrucin-Materek M, van Teijlingen ER, Barker S, Forrest Keenan K, Miedzybrodzka Z. Parenting a child with clubfoot: A qualitative study. J Ortho Trauma Nurs 2011;15(4):176-184.
9. Zaidman-Zait A, Jamieson JR. Providing web-based support for families of infants and young children with established disabilities. Infant Young Child 2007;20:11-25.
10. Morcuende JA, Dolan LA, Dietz FR, Ponseti IV. Radical reduction in the rate of extensive corrective surgery for clubfoot using the Ponseti method. Pediatrics 2004 Feb;113(2):376-380 [FREE Full text] [Medline: 14754952]
11. Desai L, Oprescu F, DiMeo A, Morcuende JA. Bracing in the treatment of children with clubfoot: past, present, and future. Iowa Orthop J 2010;30:15-23 [FREE Full text] [Medline: 21045966]
12. McKay SD, Dolan LA, Morcuende JA. Treatment results of late-relapsing idiopathic clubfoot previously treated with the Ponseti method. J Pediatr Orthop 2012 Jun;32(4):406-411. [doi: 10.1097/BPO.0b013e318256117c] [Medline: 22584843]
13. Zionts LE, Sangiorgio SN, Ebramzadeh E, Morcuende JA. The current management of idiopathic clubfoot revisited: results of a survey of the POSNA membership. J Pediatr Orthop 2012;32(5):515-520. [doi: 10.1097/BPO.0b013e318259ff79] [Medline: 22706469]
14. Masrouha KZ, Morcuende JA. Relapse after tibialis anterior tendon transfer in idiopathic clubfoot treated by the Ponseti method. J Pediatr Orthop 2012;32(1):81-84. [doi: 10.1097/BPO.0b013e31823db19d] [Medline: 22173393]
15. Halanski MA, Davison JE, Huang JC, Walker CG, Walsh SJ, Crawford HA. Ponseti method compared with surgical treatment of clubfoot: a prospective comparison. J Bone Joint Surg Am 2010 Feb;92(2):270-278. [doi: 10.2106/JBJS.H.01560] [Medline: 20124052]
16. Nemeth B. The diagnosis and management of common childhood orthopedic disorders. Curr Probl Pediatr Adolesc Health Care 2011 Jan;41(1):2-28. [doi: 10.1016/j.cppeds.2010.10.004] [Medline: 21195338]
17. Jowett CR, Morcuende JA, Ramachandran M. Management of congenital talipes equinovarus using the Ponseti method: a systematic review. J Bone Joint Surg Br 2011 Sep;93(9):1160-1164. [doi: 10.1302/0301-620X.93B9.26947] [Medline: 21911524]
18. Brashers DE. Communication and Uncertainty Management. J Commun 2001;51(3):477-497.
19. Babrow AS, Hines SC, Kasch CR. Managing uncertainty in illness explanation: An application of problematic integration theory. In: Whaley B, editor. Explaining illness: Research, theory, and strategies. Hillsdale, NJ: Erlbaum; 2000:41-67.
20. Babrow AS, Kasch CR, Ford LA. The many meanings of uncertainty in illness: toward a systematic accounting. Health Commun 1998;10(1):1-23. [doi: 10.1207/s15327027hc1001_1] [Medline: 16370987]
21. Brashers DE, Neidig JL, Reynolds NR, Haas SM. Uncertainty in illness across the HIV/AIDS trajectory. J Assoc Nurses AIDS Care 1998;9(1):66-77. [Medline: 9436169]
22. Gundersen T. 'One wants to know what a chromosome is': the internet as a coping resource when adjusting to life parenting a child with a rare genetic disorder. Sociol Health Illn 2011 Jan;33(1):81-95. [doi: 10.1111/j.1467-9566.2010.01277.x] [Medline: 20937053]
23. Thompson S, O'Hair HD. Advice-giving and the management of uncertainty for cancer survivors. Health Commun 2008 Jul;23(4):340-348. [doi: 10.1080/10410230802229712] [Medline: 18701998]
24. Braithwaite DO, Waldron VR, Finn J. Communication of social support in computer-mediated groups for people with disabilities. Health Commun 1999;11(2):123-151. [doi: 10.1207/s15327027hc1102_2] [Medline: 16370973]
25. Emmers TM, Canary DJ. The Effect of Uncertainty Reducing Strategies on Young Couples' Relational Repair and Intimacy. Commun Quart 1996;44(2):166-182.
26. Brashers DE, Neidig JL, Haas SM, Dobbs LK, Cardillo LW, Russell JA. Communication in the management of uncertainty: The case of persons living with HIV or AIDS. Commun Monogr 2000;67(1):63-84.
27. Fogel J, Albert MS, Schnabel F, Ditkoff AB, Neugut IA. Use of the Internet by Women with Breast Cancer. J Med Internet Res 2002;4(2).
28. Kinnane NA, Milne DJ. The role of the Internet in supporting and informing carers of people with cancer: a literature review. Support Care Cancer 2010 Sep;18(9):1123-1136. [doi: 10.1007/s00520-010-0863-4] [Medline: 20336326]
29. Fox S. The Social Life of Health Information. Washington, DC: Pew Research Center's Internet & American Life Project; 2011.

30. Forster DA, McLachlan HL, Rayner J, Yelland J, Gold L, Rayner S. The early postnatal period: exploring women's views, expectations and experiences of care using focus groups in Victoria, Australia. BMC Pregnancy Childbirth 2008;8:27 [FREE Full text] [doi: 10.1186/1471-2393-8-27] [Medline: 18644157]
31. Plantin L, Daneback K. Parenthood, information and support on the internet. A literature review of research on parents and professionals online. BMC Fam Pract 2009;10:34 [FREE Full text] [doi: 10.1186/1471-2296-10-34] [Medline: 19450251]
32. Madge C, O'Connor H. Parenting gone wired: empowerment of new mothers on the internet? Soc Cultur Geogr 2006;7(2):199-220.
33. Sarkadi A, Bremberg S. Socially unbiased parenting support on the Internet: a cross-sectional study of users of a large Swedish parenting website. Child Care Health Dev 2005 Jan;31(1):43-52. [doi: 10.1111/j.1365-2214.2005.00475.x] [Medline: 15658965]
34. Graber MA, Roller CM, Kaeble B. Readability levels of patient education material on the World Wide Web. J Fam Pract 1999 Jan;48(1):58-61. [Medline: 9934385]
35. McMullan M. Patients using the Internet to obtain health information: how this affects the patient-health professional relationship. Patient Educ Couns 2006 Oct;63(1-2):24-28. [doi: 10.1016/j.pec.2005.10.006] [Medline: 16406474]
36. Chen X, Siu LL. Impact of the media and the internet on oncology: survey of cancer patients and oncologists in Canada. J Clin Oncol 2001 Dec 1;19(23):4291-4297. [Medline: 11731511]
37. Schaffer R, Kuczynski K, Skinner D. Producing genetic knowledge and citizenship through the Internet: mothers, pediatric genetics, and cybermedicine. Sociol Health Illn 2008 Jan;30(1):145-159. [doi: 10.1111/j.1467-9566.2007.01042.x] [Medline: 18254838]
38. Hickson GB, Stewart DW, Altemeier WA, Perrin JM. First step in obtaining child health care: selecting a physician. Pediatrics 1988 Mar;81(3):333-338. [Medline: 3344176]
39. Cotten SR, Gupta SS. Characteristics of online and offline health information seekers and factors that discriminate between them. Soc Sci Med 2004 Nov;59(9):1795-1806. [doi: 10.1016/j.socscimed.2004.02.020] [Medline: 15312915]
40. Foster A. A nonlinear model of information-seeking behavior. J Am Soc Inf Sci Tec 2004;55(3):228-237.
41. Bragadóttir H. Computer-mediated support group intervention for parents. J Nurs Scholarsh 2008;40(1):32-38. [doi: 10.1111/j.1547-5069.2007.00203.x] [Medline: 18302589]
42. Coulson NS, Buchanan H, Aubeeluck A. Social support in cyberspace: a content analysis of communication within a Huntington's disease online support group. Patient Educ Couns 2007 Oct;68(2):173-178. [doi: 10.1016/j.pec.2007.06.002] [Medline: 17629440]
43. Namkoong K, DuBenske LL, Shaw BR, Gustafson DH, Hawkins RP, Shah DV, et al. Creating a bond between caregivers online: effect on caregivers' coping strategies. J Health Commun 2012;17(2):125-140 [FREE Full text] [doi: 10.1080/10810730.2011.585687] [Medline: 22004055]
44. Gustafson DH, Hawkins R, McTavish F, Pingree S, Chen WC, Volrathongchai K, et al. Internet-Based Interactive Support for Cancer Patients: Are Integrated Systems Better? J Commun 2008 Jun;58(2):238-257 [FREE Full text] [doi: 10.1111/j.1460-2466.2008.00383.x] [Medline: 21804645]
45. Stewart M, Letourneau N, Masuda JR, Anderson S, McGhan S. Online solutions to support needs and preferences of parents of children with asthma and allergies. J Fam Nurs 2011 Aug;17(3):357-379. [doi: 10.1177/1074840711415416] [Medline: 21813815]
46. Eysenbach G, Powell J, Kuss O, Sa ER. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. JAMA 2002;287(20):2691-2700. [Medline: 12020305]
47. Pandolfini C, Impicciatore P, Bonati M. Parents on the web: risks for quality management of cough in children. Pediatrics 2000 Jan;105(1):e1. [Medline: 10617738]
48. Neuendorf KA. The content analysis guidebook. Thousand Oaks, CA: Sage; 2002.
49. Ginossar T. Online participation: a content analysis of differences in utilization of two online cancer communities by men and women, patients and family members. Health Commun 2008;23(1):1-12. [doi: 10.1080/10410230701697100] [Medline: 18443988]
50. Winzelberg AJ, Classen C, Alpers GW, Roberts H, Koopman C, Adams RE, et al. Evaluation of an internet support group for women with primary breast cancer. Cancer 2003 Mar 1;97(5):1164-1173 [FREE Full text] [doi: 10.1002/cncr.11174] [Medline: 12599221]
51. Lincoln YS, Guba EG. Naturalistic inquiry. Beverly Hills: Sage Publications Inc; 1985.
52. Riffe D, Lacy SF, Fico GF. Analyzing media messages: Using quantitative content analysis in research. Mahwah, NJ: Erlbaum; 1998.
53. Koch-Weser S, Bradshaw YS, Gualtieri L, Gallagher SS. The Internet as a health information source: findings from the 2007 Health Information National Trends Survey and implications for health communication. J Health Commun 2010;15 Suppl 3:279-293. [doi: 10.1080/10810730.2010.522700] [Medline: 21154099]
54. Mo PK, Malik SH, Coulson NS. Gender differences in computer-mediated communication: a systematic literature review of online health-related support groups. Patient Educ Couns 2009 Apr;75(1):16-24. [doi: 10.1016/j.pec.2008.08.029] [Medline: 19233586]
55. Ellis D. A behavioural approach to information retrieval design. J Doc 1989;45:318-338.

56. Coulter A. Evidence based patient information. BMJ 1998;317(7153):225-226.
57. Tercyak KP, Mays D, DeMarco TA, Sharff MKE, Friedman S. Results of an online community needs assessment for psychoeducational interventions among partners of hereditary breast cancer previvors and survivors. J Med Internet Res 2012;14(1).
58. Bernhardt JM, Felter EM. Online pediatric information seeking among mothers of young children: results from a qualitative study using focus groups. J Med Internet Res 2004 Mar 1;6(1):e7 [FREE Full text] [doi: 10.2196/jmir.6.1.e7] [Medline: 15111273]
59. Hart A, Henwood F, Wyatt S. The role of the Internet in patient-practitioner relationships: findings from a qualitative research study. J Med Internet Res 2004 Sep 30;6(3):e36 [FREE Full text] [doi: 10.2196/jmir.6.3.e36] [Medline: 15471762]
60. Bekhuis T, Kreinacke M, Spallek H, Song M, O'Donnell JA. Using natural language processing to enable in-depth analysis of clinical messages posted to an Internet mailing list: a feasibility study. J Med Internet Res 2011;13(4):e98 [FREE Full text] [doi: 10.2196/jmir.1799] [Medline: 22112583]
61. Setoyama Y, Yamazaki Y, Namayama K. Benefits of peer support in online Japanese breast cancer communities: differences between lurkers and posters. J Med Internet Res 2011;13(4):e122 [FREE Full text] [doi: 10.2196/jmir.1696] [Medline: 22204869]
62. Kernisan LP, Sudore RL, Knight SJ. Information-seeking at a caregiving website: a qualitative analysis. J Med Internet Res 2010;12(3):e31 [FREE Full text] [doi: 10.2196/jmir.1548] [Medline: 20675292]
63. Bambina A. Online Social Support: The Interplay of Social Networks and Computer-Mediated Communication. Youngstown, NY: Cambria Press; 2007.
64. Wicks P, Massagli M, Kulkarni A, Dastani H. Use of an online community to develop patient-reported outcome instruments: the Multiple Sclerosis Treatment Adherence Questionnaire (MS-TAQ). J Med Internet Res 2011;13(1):e12 [FREE Full text] [doi: 10.2196/jmir.1687] [Medline: 21266318]
65. Swan M. Crowdsourced health research studies: an important emerging complement to clinical trials in the public health research ecosystem. J Med Internet Res 2012;14(2):e46 [FREE Full text] [doi: 10.2196/jmir.1988] [Medline: 22397809]

Edited by G Eysenbach; submitted 08.11.12; peer-reviewed by C Knapp, P Cipresso; comments to author 28.11.12; revised version received 17.12.12; accepted 20.12.12; published 22.01.13.

Please cite as:
Oprescu F, Campo S, Lowe J, Andsager J, Morcuende JA
Online Information Exchanges for Parents of Children With a Rare Health Condition: Key Findings From an Online Support Community
J Med Internet Res 2013;15(1):e16
URL: http://www.jmir.org/2013/1/e16/
doi:10.2196/jmir.2423
PMID:23470259

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org