Journal of Medical Internet Research

Journal Impact Factor (JIF) (2023): 5.8
Volume 16 (2014), Issue 5 ISSN 1438-8871 Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Original Papers

Reviews

Letter to the Editor

Original Paper

Short-Term Effects of a Web-Based Guided Self-Help Intervention for Employees With Depressive Symptoms: Randomized Controlled Trial

Anna S Geraedts[1,2,3*], MSc; Annet M Kleiboer[1,2,3*], PhD; Noortje M Wiezer[3,4*], PhD; Willem van Mechelen[2,3,5*], MD, PhD; Pim Cuijpers[1,2,3*], PhD

[1]VU University Amsterdam, Department of Clinical Psychology, Amsterdam, Netherlands

[2]EMGO+ Institute for Health and Care Research, VU University Amsterdam and VU University Medical Center Amsterdam, Amsterdam, Netherlands

[3]Body@Work Research Center Physical Activity, Work and Health, TNO-VU-VUmc, Amsterdam, Netherlands

[4]TNO, Hoofddorp, Netherlands

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

[*]all authors contributed equally

Corresponding Author:
Anna S Geraedts, MSc
VU University Amsterdam
Department of Clinical Psychology
Van der Boechorststraat 1
Amsterdam, 1081 BT
Netherlands
Phone: 31 205987451
Fax: 31 205988758
Email: a.s.geraedts@vu.nl

Abstract

Background: Depressive disorders are highly prevalent in the working population and are associated with excessive costs. The evidence for effective worker-directed interventions for employees with depressive symptoms is limited. Treating employees with depressive symptoms before sick leave via the Internet could be beneficial and cost saving.

Objective: In this study, we developed and tested the effectiveness of a Web-based guided self-help course for employees with depressive symptoms. We report on the posttreatment effectiveness of the intervention.

Methods: This study is a two-arm randomized controlled trial comparing a Web-based guided self-help course to care as usual (CAU). We recruited employees from 6 different companies via the companies' intranet and posters. The main inclusion criterion was elevated depressive symptoms as measured by a score of ≥16 on the Center for Epidemiological Studies Depression scale (CES-D). The intervention (Happy@Work) was based on problem-solving treatment and cognitive therapy and consisted of 6 weekly lessons. Participants were asked to submit their weekly assignment via the website after completion. They subsequently received feedback from a coach via the website. Self-report questionnaires on depressive symptoms (CES-D; primary outcome), anxiety measured by the Hospital Anxiety and Depression Scale (HADS), burnout measured by the Maslach Burnout Inventory (MBI), and work performance measured by the Health and Work Performance Questionnaire (HPQ; secondary outcomes) were completed at baseline and at posttreatment.

Results: A total of 231 employees were randomized to either the intervention group (n=116) or CAU (n=115).The posttreatment assessment was completed by 171 (74.0%) participants. Both the intervention and the CAU group showed significant improvements in the primary outcome of depressive symptoms, but no differences between the conditions was found (d=0.16, 95% CI –0.10 to 0.41, P=.29). Significant but small effects in favor of the intervention group were found for anxiety symptoms (d=0.16, 95% CI –0.09 to 0.42, P=.04) and exhaustion (d=0.17, 95% CI –0.09 to 0.43, P=.02).

Conclusions: This study showed that a Web-based guided self-help course for employees with depressive symptoms was not more effective in reducing depressive symptoms among employees than CAU. Large improvements in depressive symptoms in the CAU group were unforeseen and potential explanations are discussed.

(J Med Internet Res 2014;16(5):e121) doi:10.2196/jmir.3185

KEYWORDS

depression, employees, occupational therapy, Internet, prevention

Introduction

Depressive disorders are highly prevalent in the general [1-3] and working [4,5] population and lead to excessive costs [6,7]. Approximately 70%-85% of the costs are due to work absenteeism, work impairment, and loss of work productivity, which implies that companies pay the largest part of the total costs of depression [8-12]. In a recent Dutch cohort study, the total annual costs of work absenteeism due to depressive disorders were estimated at €242 million per 1 million employees, which equals €1.8 billion for the entire Dutch working population [8].

Research on the treatment of depression has been extensive [13-16] and many studies have shown positive effects of different psychotherapies (eg, [15,17,18]). Traditionally, most types of psychotherapies are delivered face-to-face in mental health care settings. However, there is increasing evidence for the effectiveness of guided self-help treatments that are delivered via the Internet [19-21]. Web-based treatments generally use the same techniques as face-to-face treatments, but patients can work through the treatment on their own in an often highly structured way. Web-based treatments may reduce costs and can increase efficiency in mental health care because of several advantages, such as high accessibility, no waiting list, and minimal contact with a professional therapist [22]. High accessibility may be of special benefit to employees because they will not lose work hours due to therapist visits outside the workplace and participation in Web-based treatments is more anonymous compared to face-to-face treatment.

The evidence for effective worker-directed interventions for employees with depressive symptoms is scarce [23]. Some research has been conducted on the treatment of employees who are absent from work (sick-listed employees) due to common mental health disorders, and the results of these studies are conflicting [23,24-27]. Previous research also shows that only a small percentage of employees with severe mental health problems seek professional treatment [28]. However, work-related aspects play an important role in the development and perpetuation of depression [23,29]. Work-related aspects, such as high job insecurity, can have a negative effect on symptom severity, and symptom severity can have a negative effect on work elements, such as job performance. Therefore, it is important to develop evidence-based worker-directed interventions for employees with depression that involve work-related aspects and the employability of the employee [23].

Recently, the Organisation for Economic Co-operation and Development (OECD) [4] published a report in which they concluded that employees with mental health problems are frequently treated when symptoms have become severe and that the work setting of the employee is not often discussed in treatment. They recommended to increase evidence-based workplace treatments for employees with mental health problems and to intervene in an earlier stage, preferably before sick leave. Other research further subscribes the importance of intervening before sick leave because it will prevent worsening of mental health problems and, as a result, it will reduce the costs of work absenteeism and loss of work productivity [29,30]. Unfortunately, there is almost no research available on the treatment of employees with mental health problems who are not on sick leave [31]. The study by Lexis and colleagues [31] showed positive results of a face-to-face problem-solving treatment for employees with a high risk for sick leave due to depressive symptoms. The promising results of this study indicate the importance of investment in intervening before sick leave. Providing such a preventive intervention via the Internet can have many advantages as previously mentioned.

The current randomized controlled trial evaluated the effectiveness of a Web-based guided self-help course for employees with depressive symptoms who were not on sick leave compared to care as usual (CAU). The intervention is aimed at reducing the employee's depressive symptoms, and we postulate that this may reduce work absenteeism and loss of work productivity as well, which will result in cost savings for the employer. The first aim of this study was to examine whether the Web-based intervention had a more positive effect on depressive symptoms compared to the CAU group. The second aim was to investigate if the intervention had positive effects on symptoms of anxiety, burnout, and work performance.

Methods

Recruitment of Participants

Participants were recruited via 6 different (international) companies in the Netherlands: 2 banking companies (company 1 and 2), 2 research institutes (company 3 and 4), 1 security company (company 5), and 1 university (company 6). Participants were recruited via banners and digital pamphlets on the companies' intranet or via posters (only in company 5). Recruitment took place between September 2011 and December 2012. Participants who were interested in taking part in the study could ask for more information about the study via email. When information was requested, one of the researchers sent an information leaflet and an informed consent form via email. The informed consent could be returned via post or email. After participants gave informed consent, they received a link to an online screening questionnaire via email. The study protocol, information leaflet, and informed consent form were approved by the Medical Ethics Committee of the VU University Medical Center (registration number 2011/2).

Participants were eligible to take part if they were 18 years of age or older, had elevated depressive symptoms as measured by a score of 16 or higher on the Center for Epidemiologic Studies Depression scale (CES-D), were not on partial or full sick leave, had access to the Internet and an email address, and

were employed by 1 of the 6 participating companies. Exclusion criteria were unstable (<1 month) medication use for depressive symptoms and having a legal labor dispute with the employer.

Procedure

Design

This study was a randomized controlled trial with 2 arms: a Web-based guided self-help course called Happy@Work and a CAU group. The full design of the study has been described in detail elsewhere [32].

Sample Size

The sample size was guided by the expected difference in the primary outcome (ie, depressive symptoms) between the 2 groups. Based on a power of 0.80, an alpha of.05, and an expected dropout percentage of 30%, we would need 100 participants in each condition to be able to show an effect-size Cohen's *d* of 0.50. Therefore, the total sample size was determined at 200.

Randomization

Randomization took place at an individual level after completion of the baseline measurement (questionnaire and clinical interview). We used stratification at 2 levels: (1) use of antidepressants and (2) receiving treatment from a psychologist or psychiatrist at study entrance. Block randomization was used with random blocks containing 4, 6, or 8 allocations. An independent researcher made the allocation schedule with a computerized random number generator and the investigators had no knowledge of the schedule. Participants were randomized into 2 groups: the intervention group or the CAU group. Participants were informed about the randomization outcome via email.

Interventions

Happy@Work

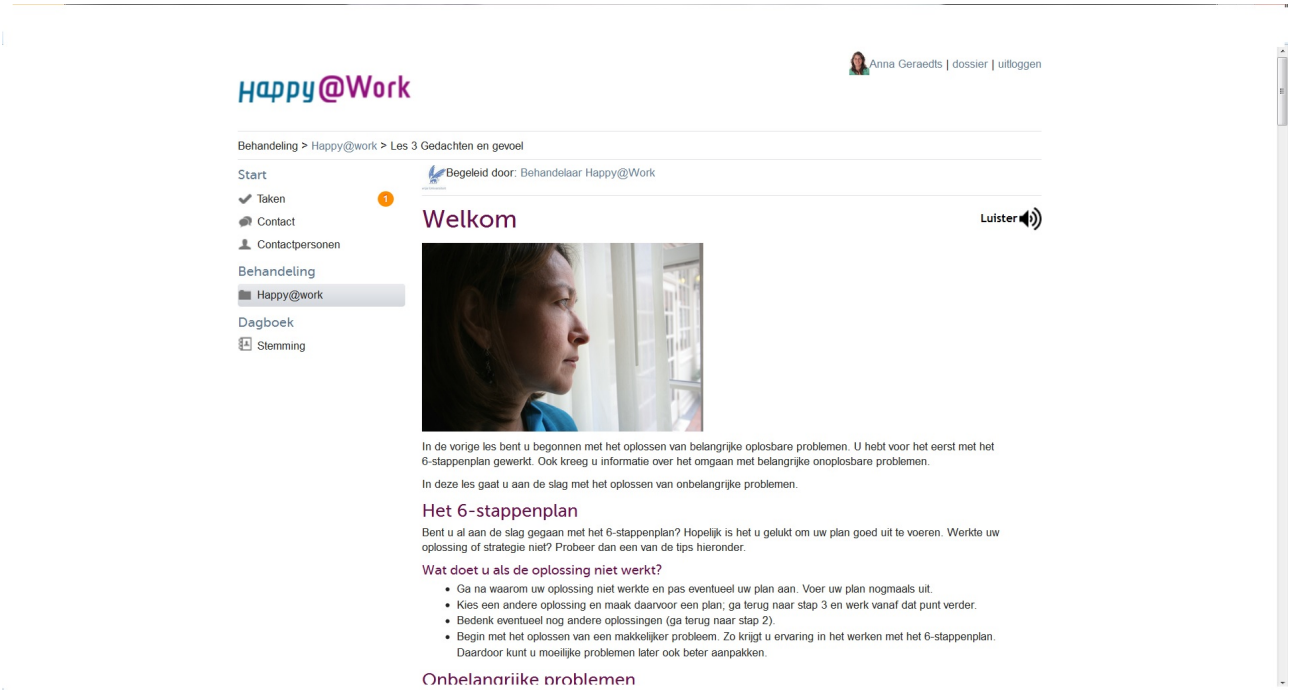
The intervention Happy@Work [33] is a brief Web-based intervention delivered with minimal guidance. It consists of 2 evidence-based treatments: problem-solving treatment (PST) [34], cognitive therapy [35], and a guideline for employees to help them to prevent work-related stress [36,37]. In PST, it is assumed that depressive symptoms can be caused by practical problems that people face in their daily lives. It is believed that when people can resolve their problems, their symptoms of depression will decrease [38]. The PST will help them solve their problems. Sometimes, however, problem solving can be disrupted by automatic thoughts such as "I am too weak to solve this problem" or "I will fail solving this problem." PST may not be sufficient to change these automatic thoughts that disrupt problem solving. Therefore, we incorporated cognitive therapy information and assignments to change these automatic thoughts in the course [35]. Some of the problems that people face are likely to be work-related. These problems are sometimes more difficult for people to comprehend [36,37]. Therefore, one lesson is focused on work-related problems specifically. Happy@Work primarily focuses on the depressive symptoms of the employee, but also incorporates psychoeducation and assignments related to dealing with stress and burnout symptoms because there is a clear relationship between the different constructs [29,39-42]. The intervention consists of 6 weekly lessons with an option of 1 week extra time in case of delay. Each lesson has a different theme, but always follows the same structure: information about the theme, examples, and assignments. Themes of the lessons are introduction of problem solving (lesson 1), problem-solving methods (lesson 2), changing cognitions (lesson 3), dealing with work-related problems (lesson 4), social support (lesson 5), and relapse prevention (lesson 6). Screenshots of the intervention can be found in Figure 1 and Multimedia Appendix 1.

At the start of the intervention, an account was generated by the researchers on the website and a coach was assigned to the participant on the website. Once the account was generated, an automatic email was sent to the participant with a link to activate the account. Participants used their email address and a self-created password to log in once the account was activated.

Participants were asked to submit their weekly assignment via the website after completion and subsequently they received feedback from a coach, again via the website, within 3 working days. Next, automatic emails were sent to participants to notify them about the feedback, to describe the following lesson, and to give the deadline for completion of the next assignment. Participants were able to start with a new lesson after they had received the feedback (ie tunneled intervention). When deadlines were not met, email reminders were sent to participants. There were no content changes, bug fixes, or periods of downtime required during the trial.

The coaches were Master's students in clinical psychology with training of 6 hours. All coaches used a protocol-treatment manual throughout the course. To ensure treatment fidelity, all feedback was reviewed by a supervisor (AG) before it was placed on the website. The support includes feedback on the assignments and motivational and empathic strategies to keep participants engaged in the course. Development of a patient-therapist alliance, as in traditional psychotherapy, was not an aim of the support.

Figure 1. Screenshot of Happy@Work intervention.



Care as Usual

Participants in the CAU group did not receive treatment or support from the coaches. However, in the email with the randomization outcome they were advised to consult their (occupational) physician or a psychologist if they wanted treatment for their depressive symptoms. Participants who were interested were sent a copy of the self-help book version of the intervention after having completed the posttreatment assessment. Participants in both conditions had access to any additional (mental) health care.

Measures

Outcome Measures

Participants filled in several questionnaires at baseline (T0) and 8 weeks later (T1). Both assessments took place online. Participants also participated in a clinical interview at T0 via telephone.

Depressive Symptoms

Symptoms of depression were measured with the Center for Epidemiological Studies Depression scale (CES-D) [43]. This questionnaire is widely used for identifying people with depressive symptoms. Its validity has been tested in different populations [44-46]. The CES-D consists of 20 items and the total score varies between 0 and 60. The Cronbach alpha in this study was .82. A score of 16 or higher represents a clinically significant level of depressive symptoms [43]. The cut-off score of 16 was used in this study as an inclusion criterion.

Anxiety Symptoms

The anxiety subscale of the Hospital Anxiety and Depression Scale (HADS) was used to measure anxiety symptoms [47]. The anxiety subscale of the HADS consists of 7 items. Scores range from 0 to 21, with higher scores indicating more anxiety. The HADS has shown good homogeneity and reliability in different normal and clinical Dutch samples [48]. The Cronbach alpha in this study was .76.

Burnout Symptoms

Burnout symptoms were measured with the Dutch version of the Maslach Burnout Inventory-General Scale (MBI) [49,50]. This self-report questionnaire contains 15 items and 3 dimensions: exhaustion (5 items), cynicism (4 items), and reduced professional efficacy (6 items). Every item was scored on a 7-point Likert scale (0-6). Following the manual of the questionnaire [50], a total score for every dimension was calculated by adding the item scores and dividing that total score by the number of items, with higher scores indicating more severe symptoms. Participants with a high score on exhaustion and a high score on cynicism or a high score on reduced professional efficacy were considered as “burnout” [50]. We rescored the professional efficacy dimension with higher scores indicating less feeling of professional efficacy, hence the high score in the burnout diagnoses. The Cronbach alphas for the different dimensions in this study were .83 for exhaustion, .83 for cynicism, and .79 for reduced professional efficacy.

Work Performance

We used the general work performance scale of the World Health Organization (WHO) Health and Work Performance Questionnaire (HPQ) [51], which contains 4 items. Item 4 gives the best and easiest indication of the subject’s perception of their own work performance [52] and we report that item only in this study. On item 4, participants were asked to rate their overall work performance during the past 4 weeks when compared to employees with comparable functions. It was scored on a 7-point Likert scale with a higher score indicating poorer work performance compared to other employees [52].

Clinical Interview

The WHO Composite International Diagnostic Interview version 2.1 (CIDI) [53] is a structured interview to assess psychiatric diagnosis defined in the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition, Text Revision (*DSM-IV-TR*) [54]. For this study, 2 sections of the CIDI were assessed: the mood disorders section and the "other" anxiety disorders (social phobia, panic disorder, agoraphobia, and generalized anxiety disorder) section. The CIDI was conducted by trained interviewers via telephone at baseline (T0) and was used for diagnostic purposes.

Health Care Utilization

A revised version of the Trimbos and iMTA Questionnaire on Costs Associated with Psychiatric Illness (TiC-P) [55] was used to collect data on health care utilization. The TiC-P is a self-report questionnaire and consists of 2 different parts that can be administrated separately. Part I was used and contains 12 items concerning health care utilization by participants. There were 2 questions added to the questionnaire on the frequency of utilization of different health care services of the company: occupational physician and occupational social work. The questionnaire was used at T0 to assess health care utilization up to 3 months before the start of the study and at posttreatment (T1) assessment to assess health care utilization between baseline and posttreatment assessment.

Other Measures

We added demographic questions, working hours, medication use for psychological problems, and treatment by a mental health specialist to the baseline questionnaire.

Statistical Analysis

Baseline Differences

Baseline differences in demographic variables and outcome measures were investigated using chi-square tests and independent-sample *t* tests.

Missing Values

Baseline data were available for all participants. Missing values at posttreatment (T1) (26.0%, 171/231) were handled by multiple imputation using the fully conditioned specified method with model type predictive mean matching in SPSS version 20.0 (IBM Corp, Armonk, NY, USA) creating 5 datasets. In multiple imputation, missing data are imputed by regression analyses and the available baseline data (demographics and baseline scores on outcome measurements) of the study completers as well as the study dropouts at posttreatment are used to estimate missing values at posttreatment. This provides a more reliable estimation of the "real" data than other imputation methods. The analyses were performed on the 5 created datasets and combined into a single overall estimate using the multiple imputation inference rules of Rubin [56]. This yields proper *P* values and confidence intervals for the estimates.

Effectiveness

Primary effectiveness analyses were conducted according to the intention-to-treat (ITT) principle. We calculated the intraclass correlation (ICC) to examine nonindependence of observations at the company level. The results showed that this was not an issue (ICC=.02, *P*=.77); therefore, a multilevel approach for analyzing the data was not deemed necessary. We performed linear regression analyses to determine differences between the intervention group and the CAU group with the posttreatment score as the dependent variable and group (intervention or CAU) as the predictor variable while controlling for baseline scores for every outcome measure. The magnitude of the effect is expressed in Cohen's *d* [57]. The Cohen's *d* was calculated by subtracting the post treatment mean score of the CAU group of the posttreatment mean score of the intervention group and dividing that result by the pooled standard deviation. Effect sizes of ≥0.8 are assumed to be large, effect sizes of 0.5-0.8 are moderate, and effect sizes of 0.2-0.5 are assumed to be small [57]. We calculated the effect sizes for all participants (ITT) and for participants who completed the intervention (per protocol analysis). Completion of the intervention was defined as completion of ≥5 lessons of the intervention. Reporting both outcomes is of high importance because intervention dropout is common in Web-based interventions.

We calculated clinical significant change as described by Jacobson and Truax [58]. This method uses a reliable change index as an index for improvement and its formula is pretest score minus posttest score divided by the standard error of the difference between the 2 test scores. To calculate the standard error of the difference between the 2 test scores in the formula, one uses the pretest SDs of the outcome and the reliability of the questionnaire. In this study, we used the following reliability scores from the questionnaires: 0.90 for the CES-D [59], 0.83 for the HADS [60], 0.88 for the exhaustion scale of MBI [50], 0.81 for the cynicism scale of MBI [50], and 0.75 for the reduced professional efficacy scale of MBI [50]. If the outcome of the sum of the reliable change index is greater than 1.96, this is considered a significant improvement because the amount of change is unlikely to have occurred by chance. The differences in clinical significant improvement rates were then expressed as odds ratios (ORs).

We calculated the recovery rates for depressive symptoms in which a score <16 on the CES-D at posttreatment was defined as recovery [59]. The differences in recovery rates were then expressed as ORs. Finally, we defined reliable recovery as clinical significant improvement between baseline and posttreatment score and recovery at posttreatment. Reliable recovery was also expressed as OR. Results are presented as the mean and standard deviations of the observed data. Between- and within-group effect sizes as well as effectiveness analyses were based on the pooled results of the imputed data.

Results

Participants and Response Rates

Overview

Figure 2 shows the flow of participants through the trial. A total of 778 employees were interested in the study and applied for more information; 320 of them did not return the informed consent form and/or did not fill in the baseline questionnaire.

We received a completed informed consent form and a baseline questionnaire from 458 participants. Of those, 208 were not eligible to take part in the study because they scored below the cut-off score of 16 on the CES-D (n=144), were already absent from work (n=48), had unstable medication use for depressive symptoms (n=12), or had a legal labor dispute with their employer (n=4). Further, 17 participants did not complete the diagnostic interview and 2 participants withdrew from the study before randomization. The remaining 231 participants were randomized to either the intervention group (n=116) or the CAU group (n=115). Most participants (n=166) were employed by 1 of the 2 banking companies, 39 by the research institutes, 11 by the security company, and 15 by the university. Of the 231 participants, 10 (4.3%) used medication without psychological treatment, 24 (10.4%) received psychological treatment but no medication, and 4 participants (1.7%) used both medication and received psychological treatment at baseline. Thus, most participants (83.6%) did not receive treatment for their depressive symptoms.

As shown in Table 1, most participants were female (62.3%, 144/231), born in the Netherlands (95.2%, 220/231), involved in an intimate relationship (76.2%, 176/231), highly educated (63.6%, 147/231), and worked for 34 hours per week on average. Participants in the intervention group were more often born outside the Netherlands (7.8%, 9/116) than participants in the CAU group (1.7%, 2/115, *P*=.03). There were no significant differences between the intervention group and the CAU group on any of the outcome measures at baseline.

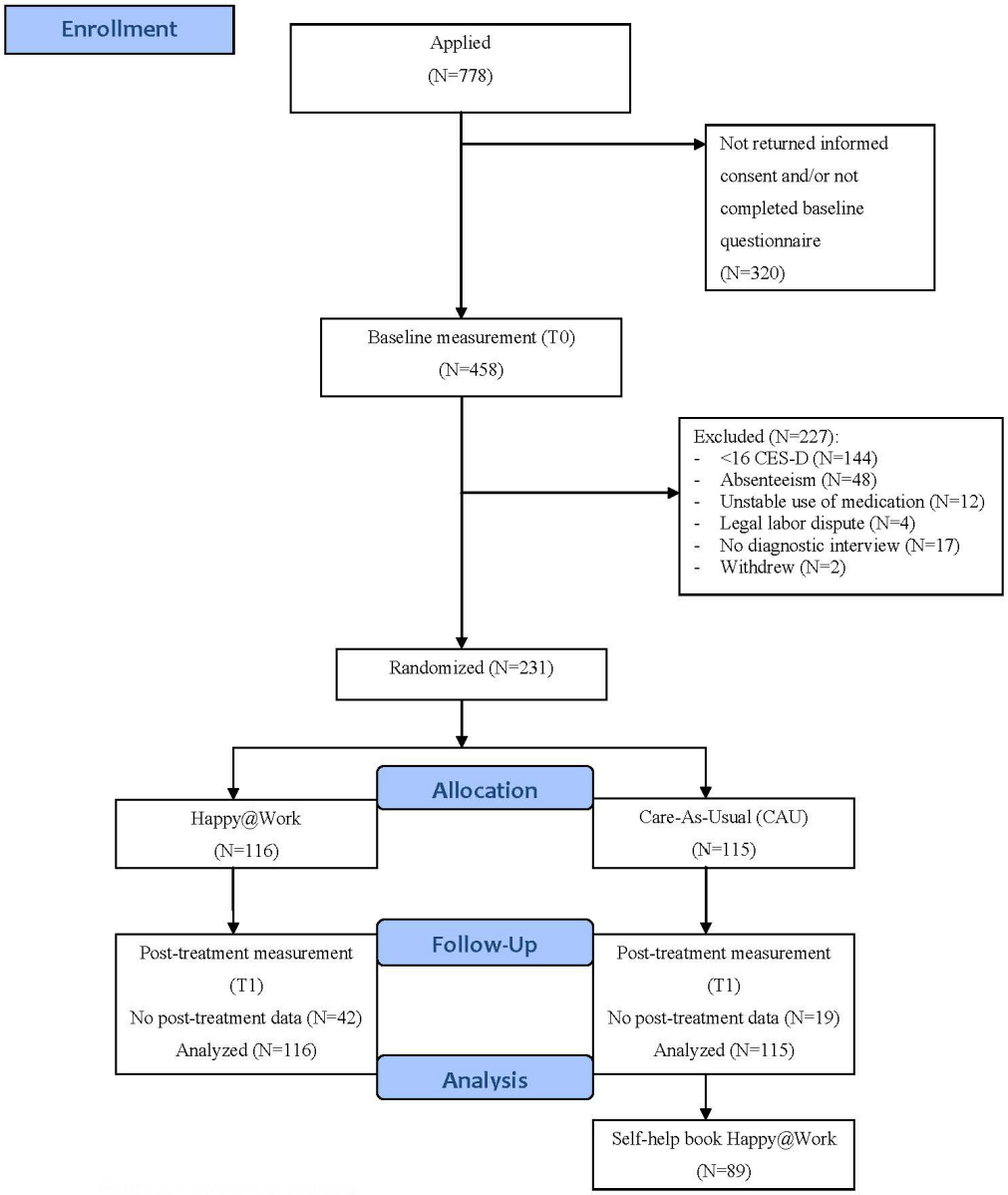
Table 1. Participants' demographic characteristics at baseline.

Characteristic	All (N=231)	Intervention (n=116)	CAU (n=115)	*P* value
Age (years), mean (SD)	43.4 (9.2)	43 (8.9)	43.8 (9.6)	.51
Gender, n (%)				.20
Female	144 (62.3)	77 (66.4)	67 (58.3)	
Male	87 (37.7)	39 (33.6)	48 (41.7)	
Country of birth, n (%)				.03
Netherlands	220 (95.2)	107(92.2)	113 (98.3)	
Other	11 (4.8)	9 (7.8)	2 (1.7)	
Marital status, n (%)				.46
Relationship	176 (76.2)	86 (74.1)	90 (78.3)	
No relationship	55 (23.8)	30 (25.9)	25 (21.7)	
Education, [a] n (%)				.25
Low	16 (6.9)	11 (9.5)	5 (4.3)	
Middle	68 (29.4)	31 (26.7)	37 (32.2)	
High	147 (63.6)	74 (63.8)	73 (63.5)	
Working hours,[b] mean (SD)	33.9 (5.0)	33.7 (4.8)	34.0 (5.3)	.65
Working days, mean (SD)	4.3 (0.7)	4.3 (0.6)	4.2 (0.7)	.32

[a]Low: primary education or lower general secondary education, middle: intermediate vocational education or high school, high: higher vocational education or university.

[b]Mean working hours per week according to contract of the employee.

Figure 2. Flowchart of participants.



Diagnosis

A total of 57 participants (24.7%) were diagnosed with a current major depressive disorder, dysthymic disorder, or both; 23 participants in the intervention group and 34 in the CAU group. Of those, 15 participants suffered from a single episode major depressive disorder (9 intervention, 6 CAU), 40 participants from a recurrent major depressive disorder (12 intervention, 28 CAU), and 9 participants from a dysthymic disorder (7 intervention, 2 CAU).

Anxiety disorders were less frequently present; a total of 48 participants (20.8%) were diagnosed, 27 participants in the intervention group and 21 in the CAU group. The different anxiety disorders frequencies were as follow: social phobia (14 intervention, 7 CAU), panic disorder without agoraphobia (2 intervention, 0 CAU), panic disorder with agoraphobia (4 intervention, 2 CAU), or generalized anxiety (11 intervention, 16 CAU) disorder.

Health Care Utilization

At posttreatment, we analyzed the health care utilization of both groups to get a more detailed view on health care utilization by the CAU group. Only a small number of the total participants made use of health care. More participants in the CAU group sought help in general health care (19 intervention, 25 CAU) or occupational health care (4 intervention, 7 CAU), but this was not statistically significant compared to the intervention group (t_{43}=0.35, P=.73; t_9=1.09, P=.31). There was only a slight difference in medication use for depressive symptoms between the groups (8 intervention, 10 CAU), and this was not statistically significant different (χ^2_1=0.00, P=.96).

Attrition and Adherence

Study Attrition

A total of 61 participants (26.4%) did not complete the T1 (posttreatment) assessment: 42/116 (36.2%) of the intervention

group and 19/115 (16.5%) of the CAU group. Data from 173 of 231 participants (74.9%) was available for the primary outcome depressive symptoms, and data from 171of 231 participants (74.0%) was available for the other outcomes. Participants in the CAU group more often completed the T1 assessment (χ^2_1=11.5, *P*=.001). Attrition rates were also lower for participants who completed the intervention (χ^2_1=32.1, *P*<.001).

Treatment Adherence

Of the participants randomized to the intervention group, 9.5% (11/116) did not start or complete the first lesson of Happy@Work. Lesson 1 was completed by 90.5% (105/116), lesson 2 by 75% (87/116), lesson 3 by 57.8% (67/116), lesson 4 by 49.1% (57/116), lesson 5 by 38.8% (45/116), and lesson 6 by 26.7% (32/116). A total 29 of 116 participants dropped out of the intervention, the other participants were not able to complete more lessons within the time limit of 7 weeks. Most participants who dropped out did not report a reason for dropout. When reasons were reported (14/116) they pertained mostly to lack of time.

Effects

Improvements on Outcome Measures

All participants improved between baseline and posttreatment on all outcomes measured (see Table 2). For the primary outcome, depressive symptoms, a high within-group effect-size Cohen's *d* was found for both the intervention (*d*=1.03, 95% CI 0.76-1.30, *P*=.001) and the CAU group (*d*=0.98, 95% CI 0.71-1.25, *P*<.001). However, there was no difference between both groups (*d*=0.16, 95% CI –0.10 to 0.41, *P*=.29). The same result was found when we compared course completers (n=45) with CAU (*d*=0.29, 95% CI –0.05 to 0.64, *P*=.13).

Small to medium within-group effects were found for the secondary outcomes anxiety, burnout, and work performance in both the intervention group and the CAU group (see Table 2). Between-group differences were small for all secondary outcomes, but participants in the intervention group improved significantly more on anxiety symptoms (*d*=0.16, 95% CI –0.09 to 0.42, *P*=.04) and the exhaustion dimension of the MBI (*d*=0.17, 95% CI –0.09 to 0.43, *P*=.02) compared to CAU. Course completers also improved more on anxiety symptoms compared to CAU (*d*=0.19, 95% CI –0.16 to 0.53, *P*=.04), but not on the exhaustion dimension of the MBI (*d*=0.17, 95% CI –0.18 to 0.52, *P*=.14). No significant between-group differences were found on the cynicism dimension of the MBI, the reduced professional efficacy dimension of the MBI, and work performance.

Table 2. Effects of intervention (n=116) versus care-as-usual (n=115) group with course completers (CC).

Outcome	Pretest, mean (SD)	Posttest, mean (SD)	Effect size,[a] Cohen's *d* (95% CI)		
			Within	Between (all)	Between (CC, n=45)
CES-D				0.16 (–0.10, 0.41)	0.29 (–0.05, 0.64)
Intervention	25.7 (7.5)	15.8 (10.6)	1.03 (0.76, 1.30)[b]		
CAU	26.1 (7.0)	18.3 (9.1)	0.98 (0.71, 1.25)[c]		
Intervention-CC	25.3 (6.5)	15.1 (10.4)			
HADS				0.16 (–0.09, 0.42)[d]	0.19 (–0.16, 0.53)[d]
Intervention	10.6 (3.8)	7.6 (3.8)	0.77 (0.51, 1.04)[c]		
CAU	10.2 (3.2)	8.3 (3.6)	0.56 (0.29, 0.82)[c]		
Intervention-CC	10.7 (3.6)	7.5 (4.0)			
MBI-exhaustion				0.17 (–0.09, 0.43)[e]	0.17 (–0.18, 0.52)
Intervention	3.3 (1.2)	2.7 (1.2)	0.50 (0.24, 0.76)[c]		
CAU	3.3 (1.1)	3.0 (1.2)	0.35 (0.09, 0.61)[c]		
Intervention-CC	3.3 (1.2)	2.7 (1.1)			
MBI-cynicism				0.30 (0.05, 0.57)	0.31 (–0.04, 0.65)
Intervention	2.8 (1.3)	2.4 (1.3)	0.31 (0.05, 0.57)[c]		
CAU	3.1 (1.3)	2.8 (1.3)	0.23 (-0.03, 0.49)[c]		
Intervention-CC	2.7 (1.2)	2.4 (1.3)			
MBI-reduced professional efficacy				0.10 (–0.16, 0.36)	0.30 (–0.05, 0.65)
Intervention	2.6 (1.0)	2.4 (1.0)	0.20 (-0.06, 0.46)[c]		
CAU	2.7 (0.9)	2.5 (0.9)	0.21 (-0.05, 0.47)[c]		
Intervention-CC	2.4 (1.0)	2.2 (1.0)			
HPQ-4				0.00 (–0.26, 0.26)	0.07 (–0.28, 0.41)
Intervention	4.1 (1.6)	3.6 (1.5)	0.32 (0.06, 0.58)[c]		
CAU	4.3 (1.8)	3.6 (1.5)	0.42 (0.16, 0.68)[c]		
Intervention-CC	4.3 (1.7)	3.4 (1.5)			

[a]The effect size is presented as Cohen's *d*: the number of standard deviations in the intervention group has improved more than the CAU group; (CAU mean–intervention mean)/pooled SD.

[b]*P*=.001.

[c]*P*<.001.

[d]*P*=.04.

[e]*P*=.02.

Clinical Significant Improvement and Reliable Recovery

Data on clinical significant improvement are reported in Table 3. Clinical significant improvement on depressive symptoms was comparable in both groups (OR 0.9, 95% CI 0.5-1.6, *P*=.82). More participants in the intervention group showed clinical significant improvement on anxiety symptoms, the exhaustion dimension of the MBI, and the cynicism dimension of the MBI compared to the CAU group, but differences between the groups were not significant (see Table 3). A total of 105 of the 231 participants (45.5%) were recovered from depression at posttreatment. More participants in the intervention group (56/116, 48.3%) recovered from depression compared to the CAU group (49/115, 42.6%), but not significantly (OR 1.3, 95% CI 0.7-2.3, *P*=.41). Reliable recovery rates for depression were also in favor for the intervention group, with an odds ratio of 1.3. In the intervention group, 44.8% (52/116) reliably recovered and 39.1% (45/115) in the CAU group. However, the difference was not statistically significant (OR 1.3, 95% CI 0.7-2.3, *P*=.39).

Table 3. Participants with clinical significant improvement.

Outcome	Intervention, n (%) (n=116)	CAU, n (%) (n=115)	OR (95% CI)	*P* value
CES-D	71 (61.2)	73 (63.5)	0.9 (0.5, 1.6)	.82
HADS	36 (31.0)	24 (20.9)	1.7 (0.9, 3.3)	.11
MBI-exhaustion	34 (29.3)	21 (18.3)	1.9 (1.0, 3.7)	.07
MBI-cynicism	16 (13.8)	13 (11.3)	1.2 (0.5, 3.1)	.67
MBI-reduced professional efficacy	10 (8.6)	12 (10.4)	0.8 (0.3, 2.2)	.61

Discussion

Principal Results

This study examined the effects of a Web-based guided self-help course, Happy@Work, on depressive symptoms (primary outcome), anxiety symptoms, 3 burnout dimensions (exhaustion, cynicism, reduced professional efficacy), and work performance (secondary outcomes) compared to CAU in employees with depressive symptoms. The study did not corroborate evidence for the effectiveness of the Web-based course compared to CAU in reducing depressive symptoms. Depressive symptoms had improved substantially in both groups at posttreatment with approximately 62% of participants showing a clinically significant improvement in both conditions. Small but significant effects in favor of the intervention group were found on 2 secondary outcomes: anxiety symptoms (*d*=.16) and the burnout dimension exhaustion (*d*=.17). However, the number of people that showed a clinically significant improvement on these measures at posttreatment did not differ between both groups. We did not find additional gains of the intervention on the other outcomes cynicism, reduced professional efficacy, and work performance.

Comparison With Prior Work

The results regarding depression are not in-line with the overall positive effects of Web-based interventions that are often found when they are compared to nonactive control groups [21,22,61]. Within-group improvement in depression in the intervention condition is comparable with other studies that examined the effects of Web-based PST in the general population [38,62]. However, these studies did not show such large improvements on depression in the control group. Furthermore, these studies showed equal improvement scores on the CES-D, but baseline scores were higher compared to this study.

Several reasons may explain the large reduction in depressive symptoms in the control group. First, an explanation could be that participants in the CAU group showed spontaneous recovery, a phenomenon which is seen frequently in patients with depression and stress [62,63]. It could be possible that spontaneous recovery is higher among those who are still at work while they experience depressive symptoms. It is known from previous research that work-related aspects are of importance in the recovery of depression [64], but to our knowledge there is no research available that subscribes that work-related aspects are related to spontaneous recovery. Second, it could be possible that our method of recruitment resulted in a selection of highly motivated employees who were willing to change. This could have let to improvement by itself.

Third, being able to function and stay at work while experiencing depressive symptoms might have had a positive influence on recovery of depressive symptoms [24,63,65]. Social support from colleagues and supervisors, social identity, regular activities, and time structure are all reported as positive effects of work on mental health [66]. Fourth, the introduction of this study in the participating companies may have had beneficial effects. A company's participation in this study gives a positive and caring signal to all employees. This may create a more open environment within the company and it might, therefore, be possible that several participants in the CAU group discussed their mental health problems with their supervisor. This may have resulted in recovery of depressive symptoms. Finally, participants in the CAU group received an email with the randomization outcome in which they were also informed about their level of depressive symptoms and they were advised to seek treatment or help for their complaints. This email could have instigated a behavioral change in such a way that participants in the CAU group moved from the preparation stage to the action stage, according to the stages-of-change model from Prochaska and colleagues [67]. Only a small percentage of the participants in the CAU group sought help in professional health care. However, it could be possible that other participants sought help in a different way; for example, via their significant other or via other self-help treatments.

Between-group significant effects were found on 2 of the secondary outcomes in this study: anxiety symptoms and exhaustion (one of the dimensions of burnout). However, these results were not confirmed by clinically significant improvement scores, which may have to do with a lack of power. It is remarkable that significant effects on anxiety symptoms but not on depressive symptoms were found because anxiety and depression often co-occur [68]. The between-group effect sizes for anxiety and depression were similar, but only the effect size on anxiety symptoms was significant. This may indicate that improvement on anxiety symptoms is more stable for all participants. Not many studies examine burnout as a secondary outcome in studies on depression treatment. However, when burnout was assessed in intervention studies with a focus on employees with stress or mental health problems, results were often conflicting. Some studies found positive results on dimensions of burnout and others have not [69,70]. This study only showed significant effect on the exhaustion dimension of burnout. This dimension is sometimes seen as the core dimension of burnout and best reflects the subject's incapacity

to act, primarily because of fatigue [71,72], which is also a symptom of depression. The other 2 dimensions, cynicism and reduced professional efficacy, reflect the subject's willingness to act, which indicates a certain attitude or cognition [72]. Therefore, it may be possible to find effects in the long term instead of directly after treatment because it is more difficult to change an attitude in a short period of time. This result also indicates the importance of a specific focus on work-related aspects, stress, and burnout symptoms within the intervention. However, interpretation of the positive results on anxiety symptoms and exhaustion needs to be done with care because these effects could have occurred by chance and the effect sizes were very small.

Limitations

This study has several limitations. The first has to do with the response rate and missing data handling. We were confronted with a high attrition rate which is seen more often in Web-based interventions [73,74]. Attrition was significantly higher in the intervention group, but we could not find baseline differences between the groups (except for country of birth) to identify possible selection bias. The bias that still may have been introduced was accounted for by applying multiple imputation techniques. Nevertheless, imputing 26% of the data may have led to unreliable estimates.

Second, completion of the intervention in this study was low compared with several other studies [38,62,75,76]. Only 26.7% of the participants completed the entire course within 7 weeks, and 38.8% completed lesson 5 within 7 weeks. Therefore, our analysis of improvement scores in the subgroup of course completers has a lack of power. If the course completion was higher, the higher effect size (d=0.29) for depressive symptoms may possibly have been significant. The low completion rate of the intervention is highly influenced by the fact that participants only received 7 weeks to complete the intervention due to the study setting. Most participants were simply not able to complete more lessons within 7 weeks and only a few participants stopped the intervention at their own request due to lack of time or other reasons. We sent several email reminders to increase completion rates, but it may have been beneficial to use other methods as well, such as telephone support in addition to Web-based support [77], increased use of persuasive technology elements [78], or tailored feedback [79]. However, it is not yet clear what methods are effective in reducing dropout of Web-based interventions [74]. Third, the participants in this study were primarily Dutch white-collar workers with high educational levels. Therefore, it is uncertain whether the results can be generalized to the general working population. It is also possible that we only recruited employees with high job security. Because of an economic recession at the time of study, recruitment job security was low for many employees. It may have been possible that employees with depressive symptoms who were experiencing low job security did not apply for this study because they were afraid for their privacy. Finally, although the use of a diagnostic interview is a great strength of this study, we did not assess the CIDI interview at posttreatment; therefore, we do not know how many patients met the criteria for major depressive disorder at posttreatment.

Implications and Future Research

The results of this study implicate that the intervention Happy@Work is not more effective in reducing depressive symptoms than CAU immediately after the intervention. All participants improved substantially between the 2 assessments on depressive symptoms and significant effects in favor of the intervention group were found on anxiety symptoms and emotional exhaustion. Several explanations may account for the high improvement rates in the CAU group. More research is needed to examine the possibilities of using e-mental health in the worksite setting and future research should further explore the needs of employees with mental health problems. Definitive conclusions about the effectiveness of the intervention can be made once long-term effects of the intervention are known.

Conclusions

This study gives a first impression of the short-term effects of a Web-based guided self-help course for employees with depressive symptoms who are not on sick leave. High within-group effect sizes for both the intervention and the CAU group were found for depressive symptoms. Statistically significant between-group effects in favor of the intervention were found for anxiety symptoms and the exhaustion dimension of the MBI, but with small effect sizes. Clinical significant improvement and reliable recovery effects did not show any significant effects in favor of the intervention. Long-term effects of this intervention need to be studied, including the role of possible mediators and moderators, as well as the cost-effectiveness of the intervention.

Acknowledgments

This study is funded by Body@Work Research Center for Physical Activity, Work and Health, TNO VUMC, Amsterdam and the EMGO Institute for Health and Care Research, VU University Amsterdam, and VU University Medical Center Amsterdam.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Screenshots from the participant view on the platform.

[PDF File (Adobe PDF File), 237KB - jmir_v16i5e121_app1.pdf]

Multimedia Appendix 2

CONSORT-EHEALTH checklist V1.6.2 [80].

[PDF File (Adobe PDF File), 991KB - jmir_v16i5e121_app2.pdf]

References

1. Alonso J, Angermeyer MC, Bernert S, Bruffaerts R, Brugha TS, Bryson H, ESEMeD/MHEDEA 2000 Investigators, European Study of the Epidemiology of Mental Disorders (ESEMeD) Project. Prevalence of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. Acta Psychiatr Scand Suppl 2004(420):21-27. [doi: 10.1111/j.1600-0047.2004.00327.x] [Medline: 15128384]
2. Kessler RC, McGonagle KA, Zhao S, Nelson CB, Hughes M, Eshleman S, et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. Arch Gen Psychiatry 1994 Jan;51(1):8-19. [Medline: 8279933]
3. Waraich P, Goldner EM, Somers JM, Hsu L. Prevalence and incidence studies of mood disorders: a systematic review of the literature. Can J Psychiatry 2004 Feb;49(2):124-138. [Medline: 15065747]
4. Organisation for Economic Co-operation and Development. OECD Publishing. Paris, France: OECD Publishing; 2012. Sick on the job? Myths and realities about mental health and work URL: http://www.oecd.org/els/emp/sickonthejob2011.htm [WebCite Cache ID 6PLEfgo0F]
5. Wang J, Adair CE, Patten SB. Mental health and related disability among workers: a population-based study. Am J Ind Med 2006 Jul;49(7):514-522. [doi: 10.1002/ajim.20319] [Medline: 16634079]
6. Berto P, D'Ilario D, Ruffo P, Di Virgilio R, Rizzo F. Depression: cost-of-illness studies in the international literature, a review. J Ment Health Policy Econ 2000 Mar 1;3(1):3-10. [Medline: 11967432]
7. Greenberg PE, Birnbaum HG. The economic burden of depression in the US: societal and patient perspectives. Expert Opin Pharmacother 2005 Mar;6(3):369-376. [doi: 10.1517/14656566.6.3.369] [Medline: 15794728]
8. de Graaf R, Tuithof M, van Dorsselaer S, ten Have M. Verzuim door psychische en somatische aandoeningen bij werkenden. Resultaten van de 'Netherlands Mental Health Survey and Incidence Study-2' (NEMESIS-2) [Absenteeism due to psychological or somatic disorders in workers. Results of the 'Netherlands Mental Health Survey and Incidence Study-2' (NEMESIS-2)]. Utrecht: Trimbos-Instituut; 2011.
9. Henderson M, Glozier N, Holland Elliott K. Long term sickness absence. BMJ 2005 Apr 9;330(7495):802-803 [FREE Full text] [doi: 10.1136/bmj.330.7495.802] [Medline: 15817531]
10. Lerner D, Henke RM. What does research tell us about depression, job performance, and work productivity? J Occup Environ Med 2008 Apr;50(4):401-410. [doi: 10.1097/JOM.0b013e31816bae50] [Medline: 18404013]
11. Smit F, Cuijpers P, Oostenbrink J, Batelaan N, de Graaf R, Beekman A. Costs of nine common mental disorders: implications for curative and preventive psychiatry. J Ment Health Policy Econ 2006 Dec;9(4):193-200. [Medline: 17200596]
12. Thomas CM, Morris S. Cost of depression among adults in England in 2000. Br J Psychiatry 2003 Dec;183:514-519 [FREE Full text] [Medline: 14645022]
13. Churchill R, Hunot V, Corney R, Knapp M, McGuire H, Tylee A, et al. A systematic review of controlled trials of the effectiveness and cost-effectiveness of brief psychological treatments for depression. Health Technol Assess 2001;5(35):1-173 [FREE Full text] [Medline: 12387733]
14. Cuijpers P, Dekker J, Hollon SD, Andersson G. Adding psychotherapy to pharmacotherapy in the treatment of depressive disorders in adults: a meta-analysis. J Clin Psychiatry 2009 Sep;70(9):1219-1229. [doi: 10.4088/JCP.09r05021] [Medline: 19818243]
15. Dobson KS. A meta-analysis of the efficacy of cognitive therapy for depression. J Consult Clin Psychol 1989 Jun;57(3):414-419. [Medline: 2738214]
16. Pinquart M, Duberstein PR, Lyness JM. Treatments for later-life depressive conditions: a meta-analytic comparison of pharmacotherapy and psychotherapy. Am J Psychiatry 2006 Sep;163(9):1493-1501. [doi: 10.1176/appi.ajp.163.9.1493] [Medline: 16946172]
17. Cuijpers P, van Straten A, Warmerdam L. Problem solving therapies for depression: a meta-analysis. Eur Psychiatry 2007 Jan;22(1):9-15. [doi: 10.1016/j.eurpsy.2006.11.001] [Medline: 17194572]
18. Cuijpers P, Geraedts AS, van Oppen P, Andersson G, Markowitz JC, van Straten A. Interpersonal psychotherapy for depression: a meta-analysis. Am J Psychiatry 2011 Jun;168(6):581-592 [FREE Full text] [doi: 10.1176/appi.ajp.2010.10101411] [Medline: 21362740]
19. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. Cogn Behav Ther 2009;38(4):196-205. [doi: 10.1080/16506070903318960] [Medline: 20183695]
20. Van't Hof E, Cuijpers P, Stein DJ. Self-help and Internet-guided interventions in depression and anxiety disorders: a systematic review of meta-analyses. CNS Spectr 2009 Feb;14(2 Suppl 3):34-40. [Medline: 19238128]
21. 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]

22. 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]
23. Nieuwenhuijsen K, Bültmann U, Neumeyer-Gromen A, Verhoeven AC, Verbeek JH, van der Feltz-Cornelis CM. Interventions to improve occupational health in depressed people. Cochrane Database Syst Rev 2008 Apr(2):CD006237. [doi: 10.1002/14651858.CD006237.pub2] [Medline: 18425942]
24. Blonk RWB, Brenninkmeijer V, Lagerveld SE, Houtman ILD. Return to work: A comparison of two cognitive behavioural interventions in cases of work-related psychological complaints among the self-employed. Work & Stress 2006 Apr;20(2):129-144. [doi: 10.1080/02678370600856615]
25. van der Klink JJ, Blonk RW, Schene AH, van Dijk FJ. Reducing long term sickness absence by an activating intervention in adjustment disorders: a cluster randomised controlled design. Occup Environ Med 2003 Jun;60(6):429-437 [FREE Full text] [Medline: 12771395]
26. Rebergen DS, Bruinvels DJ, Bezemer PD, van der Beek AJ, van Mechelen W. Guideline-based care of common mental disorders by occupational physicians (CO-OP study): a randomized controlled trial. J Occup Environ Med 2009 Mar;51(3):305-312. [doi: 10.1097/JOM.0b013e3181990d32] [Medline: 19225417]
27. Schene AH, Koeter MW, Kikkert MJ, Swinkels JA, McCrone P. Adjuvant occupational therapy for work-related major depression works: randomized trial including economic evaluation. Psychol Med 2007 Mar;37(3):351-362. [doi: 10.1017/S0033291706009366] [Medline: 17112401]
28. Hilton MF, Scheurer RW, Sheridan J, Cleary CM, Whiteford HA. Employee psychological distress and treated prevalence by indices of rurality. Aust N Z J Public Health 2010 Oct;34(5):458-465. [doi: 10.1111/j.1753-6405.2010.00590.x] [Medline: 21040172]
29. Szeto AC, Dobson KS. Mental disorders and their association with perceived work stress: an investigation of the 2010 Canadian Community Health Survey. J Occup Health Psychol 2013 Apr;18(2):191-197. [doi: 10.1037/a0031806] [Medline: 23458060]
30. Järvisalo J, Andersson B, Boedeker W, Houtman I. Mental Disorders as a Major Challenge in Prevention of Work Disability: Experiences in Finland, Germany, the Netherlands and Sweden. Helsinki: Social Insurance Institution, Research department; 2005.
31. Lexis MA, Jansen NW, Huibers MJ, van Amelsvoort LG, Berkouwer A, Tjin A Ton G, et al. Prevention of long-term sickness absence and major depression in high-risk employees: a randomised controlled trial. Occup Environ Med 2011 Jun;68(6):400-407. [doi: 10.1136/oem.2010.057877] [Medline: 20924024]
32. Geraedts AS, Kleiboer AM, Wiezer NM, van Mechelen W, Cuijpers P. Web-based guided self-help for employees with depressive symptoms (Happy@Work): design of a randomized controlled trial. BMC Psychiatry 2013;13:61 [FREE Full text] [doi: 10.1186/1471-244X-13-61] [Medline: 23418886]
33. Happy@Work. URL: https://happy-at-work.e-behandeling.nl/ [accessed 2013-12-13] [WebCite Cache ID 6Lq0BElgt]
34. Bowman D, Scogin F, Lyrene B. The efficacy of self-examination therapy and cognitive bibliotherapy in the treatment of mild to moderate depression. Psychotherapy Research 1995 Jan;5(2):131-140. [doi: 10.1080/10503309512331331256]
35. Beck AT, Rush AJ, Shaw BF, Emery G. Cognitive Therapy of Depression. New York: Guilford Press; 1979.
36. Franck E, Wiezer N. Handleiding Preventie Leidraad [Manual for Prevention Guideline]. Hoofddorp: TNO; 2004.
37. Franck E, Wiezer N. Ervaringen met de Preventie Leidraad [Experiences with the Prevention Guideline]. Hoofddorp: TNO; 2004.
38. Warmerdam L, van Straten A, Twisk J, Riper H, Cuijpers P. Internet-based treatment for adults with depressive symptoms: randomized controlled trial. J Med Internet Res 2008;10(4):e44 [FREE Full text] [doi: 10.2196/jmir.1094] [Medline: 19033149]
39. Brenninkmeyer V, Van Yperen NW, Buunk BP. Burnout and depression are not identical twins: is decline of superiority a distinguishing feature? Personality and Individual Differences 2001 Apr;30(5):873-880. [doi: 10.1016/S0191-8869(00)00079-9]
40. Hakanen JJ, Schaufeli WB. Do burnout and work engagement predict depressive symptoms and life satisfaction? A three-wave seven-year prospective study. J Affect Disord 2012 Dec 10;141(2-3):415-424. [doi: 10.1016/j.jad.2012.02.043] [Medline: 22445702]
41. Hammen C. Stress and depression. Annu Rev Clin Psychol 2005;1:293-319. [doi: 10.1146/annurev.clinpsy.1.102803.143938] [Medline: 17716090]
42. Maslach C, Schaufeli WB, Leiter MP. Job burnout. Annu Rev Psychol 2001 Feb;52:397-422. [doi: 10.1146/annurev.psych.52.1.397] [Medline: 11148311]
43. Bouma J, Ranchor AV, Sanderman R, van Sonderen E. Het meten van symptomen van depressie met de CES-D; een handleiding [Assessment of Symptoms of Depression by Means of the CES-D: A Manual]. Groningen: Stichting Drukkerij Regenboog; 1995.
44. Furukawa T, Hirai T, Kitamura T, Takahashi K. Application of the Center for Epidemiologic Studies Depression Scale among first-visit psychiatric patients: a new approach to improve its performance. J Affect Disord 1997 Oct;46(1):1-13. [Medline: 9387082]

45. Haringsma R, Engels GI, Beekman AT, Spinhoven P. The criterion validity of the Center for Epidemiological Studies Depression Scale (CES-D) in a sample of self-referred elders with depressive symptomatology. Int J Geriatr Psychiatry 2004 Jun;19(6):558-563. [doi: 10.1002/gps.1130] [Medline: 15211536]
46. Yang HJ, Soong WT, Kuo PH, Chang HL, Chen WJ. Using the CES-D in a two-phase survey for depressive disorders among nonreferred adolescents in Taipei: a stratum-specific likelihood ratio analysis. J Affect Disord 2004 Nov 1;82(3):419-430. [doi: 10.1016/j.jad.2004.04.008] [Medline: 15555693]
47. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983 Jun;67(6):361-370. [Medline: 6880820]
48. Spinhoven P, Ormel J, Sloekers PP, Kempen GI, Speckens AE, Van Hemert AM. A validation study of the Hospital Anxiety and Depression Scale (HADS) in different groups of Dutch subjects. Psychol Med 1997 Mar;27(2):363-370. [Medline: 9089829]
49. Maslach C, Jackson SE, Leiter MP. Maslach Burnout Inventory Manual. Palo Alto: Consulting Psychologists Press; 1996.
50. Schaufeli WB, van Dierendonck D. UBOS; Utrechtse Burnout Schaal [UBOS; Utrecht Burnout Scale]. Lisse: Swets Test Publishers; 2000.
51. Kessler RC, Barber C, Beck A, Berglund P, Cleary PD, McKenas D, et al. The World Health Organization Health and Work Performance Questionnaire (HPQ). J Occup Environ Med 2003 Feb;45(2):156-174. [Medline: 12625231]
52. Kessler RC, Ames M, Hymel PA, Loeppke R, McKenas DK, Richling DE, et al. Using the World Health Organization Health and Work Performance Questionnaire (HPQ) to evaluate the indirect workplace costs of illness. J Occup Environ Med 2004 Jun;46(6 Suppl):S23-S37. [Medline: 15194893]
53. World Health Organization. Composite International Diagnostic Interview (version 2.1). Geneva: WHO; 1990.
54. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (Text Revision). Washington: Jaypee Brothers Medical Publishers (P) Ltd; 2000.
55. Hakkaart-van Roijen L, van Straten A, Donker M, Tiemans B. Manual Trimbos/iMTA questionnaire for costs associated with psychiatric illness (TIC-P). Rotterdam: Institute for Medical Technology Assessment; 2002.
56. Rubin DB. Multiple Imputation for Nonresponse in Surveys. New York: Wiley; 1987.
57. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Hillsdale, NJ: L Erlbaum Associates; 1988.
58. Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. J Consult Clin Psychol 1991 Feb;59(1):12-19. [Medline: 2002127]
59. Radloff LS. The CES-D Scale: A self-report depression scale for research in the general population. Applied Psychological Measurement 1977 Jun 01;1(3):385-401. [doi: 10.1177/014662167700100306]
60. Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale. An updated literature review. J Psychosom Res 2002 Feb;52(2):69-77. [Medline: 11832252]
61. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. Cogn Behav Ther 2009;38(4):196-205. [doi: 10.1080/16506070903318960] [Medline: 20183695]
62. 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]
63. Zetterqvist K, Maanmies J, Ström L, Andersson G. Randomized controlled trial of internet-based stress management. Cogn Behav Ther 2003;32(3):151-160. [doi: 10.1080/16506070302316] [Medline: 16291546]
64. Lagerveld SE, Bültmann U, Franche RL, van Dijk FJ, Vlasveld MC, van der Feltz-Cornelis CM, et al. Factors associated with work participation and work functioning in depressed workers: a systematic review. J Occup Rehabil 2010 Sep;20(3):275-292 [FREE Full text] [doi: 10.1007/s10926-009-9224-x] [Medline: 20091105]
65. Lagerveld SE, Blonk RW, Brenninkmeijer V, Wijngaards-de Meij L, Schaufeli WB. Work-focused treatment of common mental disorders and return to work: a comparative outcome study. J Occup Health Psychol 2012 Apr;17(2):220-234. [doi: 10.1037/a0027049] [Medline: 22308965]
66. Harnois G, Gabriel P. Work and Health: Impact, Issues and Good Practices. Geneva: World Health Organization; 2000.
67. Prochaska JO, DiClemente CC, Norcross JC. In search of how people change. Applications to addictive behaviors. Am Psychol 1992 Sep;47(9):1102-1114. [Medline: 1329589]
68. Pollack MH. Comorbid anxiety and depression. J Clin Psychiatry 2005;66 Suppl 8:22-29. [Medline: 16336033]
69. Van Rhenen W, Blonk RW, van der Klink JJ, van Dijk FJ, Schaufeli WB. The effect of a cognitive and a physical stress-reducing programme on psychological complaints. Int Arch Occup Environ Health 2005 Mar;78(2):139-148. [doi: 10.1007/s00420-004-0566-6] [Medline: 15761748]
70. Ruwaard J, Lange A, Bouwman M, Broeksteeg J, Schrieken B. E-mailed standardized cognitive behavioural treatment of work-related stress: a randomized controlled trial. Cogn Behav Ther 2007;36(3):179-192. [doi: 10.1080/16506070701381863] [Medline: 17852171]
71. Schaufeli WB, Taris TW. The conceptualization and measurement of burnout: Common ground and worlds apart. Work & Stress 2005 Jul;19(3):256-262. [doi: 10.1080/02678370500385913]
72. Schaufeli W. Burn-out in discussie: de stand van zaken [Current debates about burnout: a review]. De Psycholoog 2007;42(10):534-540.

73. 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]
74. Melville KM, Casey LM, Kavanagh DJ. Dropout from Internet-based treatment for psychological disorders. Br J Clin Psychol 2010 Nov;49(Pt 4):455-471. [doi: 10.1348/014466509X472138] [Medline: 19799804]
75. 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]
76. 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]
77. Carlbring P, Gunnarsdóttir M, Hedensjö L, Andersson G, Ekselius L, Furmark T. Treatment of social phobia: randomised trial of internet-delivered cognitive-behavioural therapy with telephone support. Br J Psychiatry 2007 Feb;190:123-128 [FREE Full text] [doi: 10.1192/bjp.bp.105.020107] [Medline: 17267928]
78. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
79. Carlbring P, Maurin L, Törngren C, Linna E, Eriksson T, Sparthan E, et al. Individually-tailored, Internet-based treatment for anxiety disorders: A randomized controlled trial. Behav Res Ther 2011 Jan;49(1):18-24. [doi: 10.1016/j.brat.2010.10.002] [Medline: 21047620]
80. 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

CAU: care as usual
CES-D: Center for Epidemiological Studies Depression scale
CIDI: Composite International Diagnostic Interview
HADS: Hospital Anxiety and Depression Scale
HPQ: Health and Work Performance Questionnaire
ICC: intraclass correlation
ITT: intention-to-treat
MBI: Maslach Burnout Inventory-General Scale
PST: problem-solving treatment

Edited by G Eysenbach; submitted 13.12.13; peer-reviewed by C Spates, J Hoof, S Rice; comments to author 13.03.14; revised version received 20.03.14; accepted 18.04.14; published 06.05.14.

Please cite as:
Geraedts AS, Kleiboer AM, Wiezer NM, van Mechelen W, Cuijpers P
Short-Term Effects of a Web-Based Guided Self-Help Intervention for Employees With Depressive Symptoms: Randomized Controlled Trial
J Med Internet Res 2014;16(5):e121
URL: http://www.jmir.org/2014/5/e121/
doi:10.2196/jmir.3185
PMID:24800966

Review

Computer-Delivered and Web-Based Interventions to Improve Depression, Anxiety, and Psychological Well-Being of University Students: A Systematic Review and Meta-Analysis

E Bethan Davies[1], BSc (Hons), MSc (Hons); Richard Morriss[1], MBChB, MMedSci, MD, FRCPsych; Cris Glazebrook[1], RGN, PhD, CPsychol

Division of Psychiatry and Applied Psychology, School of Medicine, Institute of Mental Health, The University of Nottingham, Nottingham, United Kingdom

Corresponding Author:
E Bethan Davies, BSc (Hons), MSc (Hons)
Division of Psychiatry and Applied Psychology, School of Medicine
Institute of Mental Health
The University of Nottingham
Jubilee Campus, Triumph Road
Nottingham, NG7 2TU
United Kingdom
Phone: 44 0115 74 84293
Fax: 44 0115 823 1289
Email: mcxebd@nottingham.ac.uk

Abstract

Background: Depression and anxiety are common mental health difficulties experienced by university students and can impair academic and social functioning. Students are limited in seeking help from professionals. As university students are highly connected to digital technologies, Web-based and computer-delivered interventions could be used to improve students' mental health. The effectiveness of these intervention types requires investigation to identify whether these are viable prevention strategies for university students.

Objective: The intent of the study was to systematically review and analyze trials of Web-based and computer-delivered interventions to improve depression, anxiety, psychological distress, and stress in university students.

Methods: Several databases were searched using keywords relating to higher education students, mental health, and eHealth interventions. The eligibility criteria for studies included in the review were: (1) the study aimed to improve symptoms relating to depression, anxiety, psychological distress, and stress, (2) the study involved computer-delivered or Web-based interventions accessed via computer, laptop, or tablet, (3) the study was a randomized controlled trial, and (4) the study was trialed on higher education students. Trials were reviewed and outcome data analyzed through random effects meta-analyses for each outcome and each type of trial arm comparison. Cochrane Collaboration risk of bias tool was used to assess study quality.

Results: A total of 17 trials were identified, in which seven were the same three interventions on separate samples; 14 reported sufficient information for meta-analysis. The majority (n=13) were website-delivered and nine interventions were based on cognitive behavioral therapy (CBT). A total of 1795 participants were randomized and 1480 analyzed. Risk of bias was considered moderate, as many publications did not sufficiently report their methods and seven explicitly conducted completers' analyses. In comparison to the inactive control, sensitivity meta-analyses supported intervention in improving anxiety (pooled standardized mean difference [SMD] −0.56; 95% CI −0.77 to −0.35, $P<.001$), depression (pooled SMD −0.43; 95% CI −0.63 to −0.22, $P<.001$), and stress (pooled SMD −0.73; 95% CI −1.27 to −0.19, $P=.008$). In comparison to active controls, sensitivity analyses did not support either condition for anxiety (pooled SMD −0.18; 95% CI −0.98 to 0.62, $P=.66$) or depression (pooled SMD −0.28; 95% CI −0.75 to −0.20, $P=.25$). In contrast to a comparison intervention, neither condition was supported in sensitivity analyses for anxiety (pooled SMD −0.10; 95% CI −0.39 to 0.18, $P=.48$) or depression (pooled SMD −0.33; 95% CI −0.43 to 1.09, $P=.40$).

Conclusions: The findings suggest Web-based and computer-delivered interventions can be effective in improving students' depression, anxiety, and stress outcomes when compared to inactive controls, but some caution is needed when compared to other trial arms and methodological issues were noticeable. Interventions need to be trialed on more heterogeneous student samples

and would benefit from user evaluation. Future trials should address methodological considerations to improve reporting of trial quality and address post-intervention skewed data.

(J Med Internet Res 2014;16(5):e130) doi:10.2196/jmir.3142

KEYWORDS

systematic review; meta-analysis; intervention; universities; students; mental health; depression; anxiety; health promotion

Introduction

Depression and anxiety are common mental health problems experienced by university students [1]. A recent review reported a 30.6% mean prevalence rate of depression in students [2] and a cross-sectional survey reported 17.3% prevalence of clinically-significant psychiatric caseness in a UK student sample [3]. Being in higher education is associated with many stressors and transitional events, and students fall within the age range when common mental health problems are at their developmental peak [4]. Of students who screened below the threshold for anxiety and depression at entry to university, 9% were above the threshold for depression and 20% for anxiety 18 months into their course [5]. Depression and anxiety can impair students' academic performance and social functioning, cause significant burden at university, and potentially affect their future career opportunities [4,6,7]. Students' help-seeking behavior for their mental health difficulties is limited, with many not contacting relevant professional services [8]. Young people do not seek out help for several reasons, including personal preferences for self-reliance in managing their mental health [9].

Computer-delivered and Internet-enabled interventions have been increasingly trialed in recent years [10]. Programming technology means interventions can be delivered using a range of multimedia formats and interactive features to engage users and facilitate intervention efficacy [11]. Computer- and Internet-delivered interventions hold many advantages; they can be tailored to student needs, accessed anonymously, and provide a more comfortable private environment to access sensitive information [12]. Online interventions can be a form of outreach to individuals who may not access traditional face-to-face services [13]. Evidence-based psychotherapies have been effectively adapted for Internet-based delivery, with much evidence supporting computer-delivered cognitive behavioral therapy (CCBT) in improving depression and/or anxiety outcomes [14-17]. The Internet is an essential tool for higher education and thus highly accessible to students [12,18]. Students also use the Internet for health-related purposes; over a third of students stated that information found via the Internet had a significant effect on their own health self-care [18]. Given that students may not seek out professional help for their mental health, computerized technologies could provide access to self-help. Students may have favorable preferences toward self-help due to their increasing independence and ability to be self-reliant during their transition to young adulthood [19]. Over half of students in an Australian sample who screened for high psychological distress reported strong intentions in using an online program for student well-being [12]. As Internet-based interventions have been cited as an approach that may be particularly engaging and useful for higher education students given their limited help-seeking behavior [12,20-22], there is a need to identify and synthesize the evidence from these types of interventions for improving common mental health difficulties in higher education populations. Several UK universities appear to offer online counselling to their students, but students still have to engage in help-seeking behavior to access these services and may have stigmatizing attitudes toward professional help [23]. Self-guided computer and Internet-based resources may help to avoid this stigma and be in line with preferences for self-reliance. The recent systematic review by Farrer and colleagues [4] explored technology-based interventions trialed in higher education populations and has provided a comprehensive narrative appraisal of these trials. However, quantitative analysis was not conducted due to the variation of technologies employed in the studies. We hope to expand on this by focusing only on interventions delivered through websites and offline computer programs for improving mental health outcomes, and conducting meta-analysis to explore these outcomes. Analysis of this type of intervention in student populations has not been explored previously. The aim of this review is to explore whether computer-delivered and Web-based (ie, website-based) interventions are effective in improving depression, anxiety, and psychological well-being in higher education students.

Methods

Search Methodology and Identification of Trials

Nine electronic databases, including PsychINFO, CENTRAL, and PsychMed, were searched in March-April 2012; the search was repeated in June 2013 to ensure the search was as current as possible. Search terms (Multimedia Appendix 1) were developed through literature review and related to Internet- and computer-delivered interventions, mental health, and higher education. Several publisher websites, published reviews, and intervention studies were hand-searched. There was no restriction in year or language of publications. Studies met the following eligibility criteria:

1. The interventions had to aim to improve psychological distress, stress, depressive, or anxiety symptomology, and had administered valid and reliable measure(s) reflecting this symptomology. Interventions that also addressed general aspects of psychological well-being (eg, sleep) and included a primary mental health outcome were also included.
2. The intervention was delivered via a website or offline computer program and accessed via computer, laptop, or other technological device (eg, tablet). These technological mediums were used as a medium for delivering the intervention. Human support was included in the review

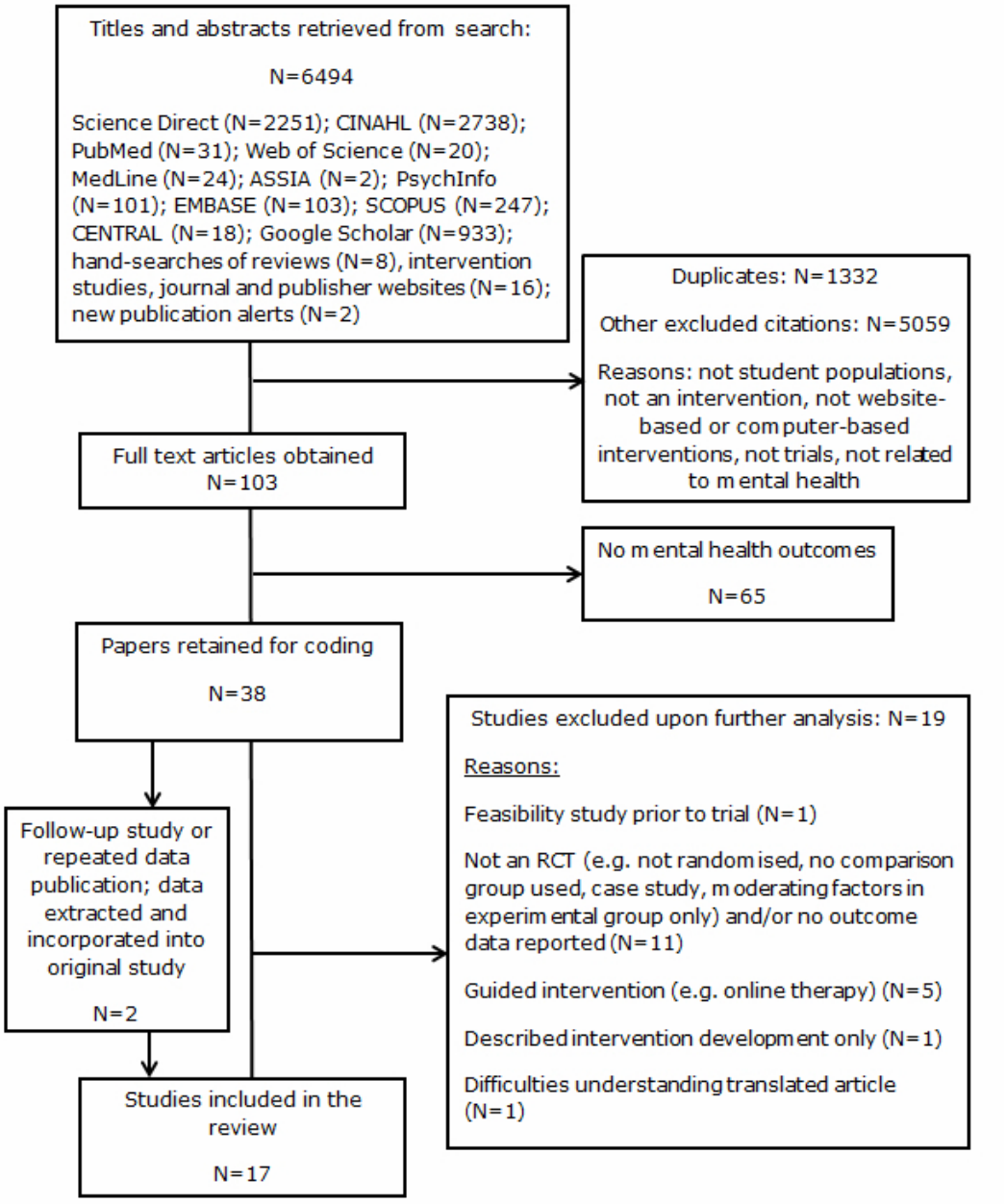
providing it was delivered by laypersons or non-health care professionals and was a complementary component of intervention.

3. The study was published in a peer-reviewed journal.
4. The intervention was trialed through randomized controlled trial (RCT) design. Trial arms need to consist of an experimental condition and an inactive control (ie, no-treatment or wait list control) condition and/or an active control and/or comparison intervention. Active control was defined as participants who received materials designed to mimic the time and attention received by participants assigned to the intervention. Active controls were not designed to produce the same changes upon outcomes as expected in the intervention.
5. The intervention was trialed on undergraduate and/or postgraduate students in higher education institutions [HEIs]. HEIs were tertiary educational institutions, such as universities and colleges.

Secondary outcomes of interest were help-seeking behavior, mental health service utilization, diagnosis of mental disorder, and participant attrition. Interventions were excluded if there was face-to-face human support adjunct to intervention, they were not Web-based or offline computer programs, they were online support groups, or were mobile or tablet applications. Interventions that utilized computers/Internet to facilitate communication (eg, email, online counselling) between health professionals and users were also excluded as we wanted to explore whether computer-delivered and Web-based interventions were comparable to traditional therapies (eg, face-to-face CBT) and had any effects on mental health outcomes in comparison to receiving no treatment. Mobile applications (“apps”) were also excluded as, at the time of conducting the search, it was felt these were relatively new mediums in terms of therapeutic interventions and appeared more likely to be used as a device to display information in the same way as a DVD/video. Online interventions for eating disorders and alcohol/substance use were not included as these have been previously reviewed in students [24,25]. Publications were excluded if they focused on mediating effects upon outcome measures only within experimental groups, or if both the intervention and active control/comparison intervention received the same intervention materials and there was no inactive control condition.

A total of 6494 titles and abstracts were retrieved from the search and screened by EBD to address their inclusion eligibility. Reference lists of relevant reviews were also searched. The updated search resulted in inclusion of some additional studies that were not published at the time of the first search. The full text of 103 articles was obtained for further analysis and coding. Of these, 38 addressed the targeted mental health criteria and 19 were excluded as they did not meet eligibility criteria or presented translation difficulties [26] (see Multimedia Appendix 2 for further description). A total of 19 articles met inclusion criteria, which included one follow-up to an included study [27] and two publications reporting the same trial [28,29]; data from both were extracted and collapsed into the original studies, resulting in 17 citations. Figure 1 outlines the search process (also see Multimedia Appendix 3).

Figure 1. PRISMA flowchart outlining process for systematic review/meta-analysis.



Data Extraction and Assessment

Data extraction was performed by EBD using a template based on the Cochrane Review template [30] and the CONSORT checklist for reporting eHealth interventions [31]. Authors were contacted if necessary to clarify information. Data regarding post-intervention means and standard deviations from relevant mental health outcome measures and information about participant attrition were extracted from the included studies and entered into Review Manager ("RevMan") software [32].

Interventions were classified by their type of prevention [33]: "universal interventions" target a whole population regardless of individual risk and do not involve screening; "selective interventions" select individuals at some risk of a mental health disorder but without screening of mental health symptomology; "indicated interventions" target those who screen for some level of mental health symptomology but do not have a diagnosis; and "treatment interventions" are delivered to individuals with a diagnosed mental disorder [4]. For this review, "selective" and "indicated" interventions were collapsed into one category as it can be difficult to decipher whether interventions discretely fit into one category.

The level of human support provided to participants was coded using categories used previously [4,34]. Only three categories were used as we did not aim to explore interventions that involved extensive contact time between participants and a human contact. The three categories were: (1) no-contact intervention (no human face-to-face or verbal contact for any aspect of study; email contact only with participants), (2) self-administered intervention (human contact for administration of measures only), and (3) semi-guided intervention (human contact ≤90 minutes for prompts or reminders, guidance on intervention use, and/or support in completing intervention).

The Cochrane Collaboration risk of bias tool [35] was used to assess trial quality. The tool provides a checklist to aid understanding of trial quality and does not calculate an overall quality score. The tool assesses study bias across five methodological domains: sequence generation, allocation concealment, blinding, incomplete outcome data, and selective reporting.

Process for Meta-Analysis

Meta-analyses were planned to explore the effects of interventions upon depression, anxiety, stress, and psychological distress related outcomes. These outcomes were analyzed in three subgroups: (1) comparing intervention to inactive control, (2) comparing intervention to active control, and (3) intervention compared to comparison intervention. If trials conducted three or more trial arms, the trial arms were separated corresponding to the three comparison analyses. In studies using two or more active control or comparison intervention conditions, the least active control was entered into analysis. Secondary analyses were conducted to explore year of publication and use of participant incentives upon outcomes, as well as exploring rates of attrition between trial arms. Continuous data on clinical outcomes are often not normally distributed and extracted data were explored for normality via presence of skew. This is done by multiplying the standard deviation by two; if the mean is smaller than this number, it suggests the data is skewed [36]. RevMan was used for calculating effect sizes and conducting meta-analyses. Standard errors were transformed into standard deviations by multiplying the standard error by the square root of the sample size. If insufficient outcome data were reported for extraction, those studies were not included in meta-analysis. If studies reported more than one type of outcome measure for specific outcomes of interest, the measure most aligned to DSM-IV criteria for depressive and anxiety disorders was selected for analysis. The Standardized Mean Difference (SMD) is a version of effect size typically calculated in reviews and is expressed as Hedges' *g*. SMDs were calculated for each included study by subtracting the post-intervention mean of the intervention condition from the post-intervention mean of the comparison condition, and dividing this by the pooled standard deviation from both conditions [37]. Use of SMD allows for comparisons across included studies where they used different psychometric measures to assess the same outcomes [38]. Inferences of Hedges' *g* can be made using Cohen's *d* conventions as small (0.2), medium (0.5), and large (0.8) [39].

We anticipated included studies would be heterogeneous due to the different types of preventative intervention and so would differ on the baseline symptomology of participants. To help account for the expected heterogeneity, Random Effects Models (RAM) with 95% confidence intervals (CI) were applied throughout analysis. RAM assumes that included studies are trialed on different populations and so are calculating different intervention effects [38,40]. The I^2statistic was calculated to explore heterogeneity and is expressed as a percentage indicating its degree: 25% indicates low heterogeneity, 50% suggests moderate, and 75% is a threshold marker for high heterogeneity [41]. The *Q* statistic was also calculated and provides the statistical significance of heterogeneity.

Results

Intervention Characteristics

The search yielded 17 studies. The symptomology measured within trials were depression [28,42-52], anxiety [28,42-48,50], stress [46,53-55], psychological distress [50,54,56], social anxiety [52], and examination anxiety [57]. Some interventions focused on general psychological well-being: improving relationship functioning [43,44], decreasing elevated levels of perfectionism [28,42], increasing students' use of mindfulness [54], improving international students' social support, acculturation, and hardiness [56], and increasing use of lucid dreaming to help alleviate depression [51]. Of the studies, seven trials were of three interventions conducted on separate samples; therefore, there are 14 distinct interventions for review. Multimedia Appendix 4 provides a summary of included interventions.

A total of 11 trials were selective or indicated interventions, where participants were included if they were screened for specific aspects of mental health symptomology or other psychological factors [28,42,45-50,53,55,57]. Inclusion criteria included: elevated perfectionism [28,42], elevated stress [53,55], minimal/mild symptoms of depression and anxiety [45,50], low/moderate psychological distress [47], elevated anxiety sensitivity [48], elevated psychological distress [49], self-reported examination anxiety [57], and mild/moderate levels of depression, anxiety, or stress [46]. Five trials were universal, in which mental symptomology were not explicit inclusion criteria; participants had to be in ≥4 month long romantic relationships [43,44], be Indian international students [56], have no lucid dreaming experience [51], or have access to an Internet-connected computer [54]. One intervention was treatment as participants met DSM-IV diagnostic criteria for social anxiety [52]. It is difficult to decipher whether some included trials discretely fitted the selective or indicated type. Some studies recruited participants with minimal symptomology or focused on other risk factors for depression and anxiety, such as elevated perfectionism [28,42].

Of the studies, 11 contained two trial arms [42,44-46,48,49,51,54-57], with five using three arms [28,43,47,52,53], and one study with four arms [50]. Five trials compared intervention to inactive (ie, no treatment or waitlist control) and either an active control [53] or comparison intervention [28,47,50,52], five trialed the intervention to an active control [44,45,55-57], six trialed against inactive control [42,46,48,49,51,54], and one compared intervention to a comparison intervention and active control [43]. Further, 13 studies [28,42-50,52,55,57] trialed interventions based on CBT; this included seven studies in which three interventions were trialed on separate samples [28,42-44,47,49,50]. Other interventions were based on mindfulness [54], stress management theory and cognitive learning theory [53], and lucid dreaming [51].

Location and Delivery of Intervention

The majority of interventions were delivered via a website or university intranet (n=13) with four being offline computer programs [43-45,55]. Five trials were delivered at a study site,

eg, researcher-monitored computer lab [43-45,47,50,55], while participants in six Internet-based interventions accessed the intervention in their own location [48,49,52,53,56,57]. A total of 14 trials had interventions with an modular/sectional format [28,42-48,50,52,54-57] ranging from three [56] to 13 modules [28,42]. The other trials coupled module-based ("MoodGym") and psycho-educational ("BluePages") websites [49], provided biweekly instruction via a website [51], and included a psycho-educational stress management website [53]. The intervention delivery period ranged from 2 [53,54] to 12 weeks [42], with median length of 6 weeks. All studies reported short-term outcomes (≤12 weeks) with measures usually administered at the end of the trial. Five reported additional follow-up at 6 months [46,48,53], 10 months [44], and 1 year post-baseline [52]. Four Web-based interventions stated how much time was required to spend accessing the intervention: at least four 20-minute periods over 2 weeks [53]; 1 hour per week over 3 weeks [47]; 30 minutes per week over 6 weeks [57]; and 5-7 days for each module [48]. The four computer-delivered interventions took between 30 to 120 minutes [43-45,55] to complete and were supplemented by weekly standardized emails.

Use of Human Support in Interventions

Seven trials were classified as self-administered [28,42,48,51-53,57], with nine being semi-guided [43-47,50,54-56]. Participants in one trial received no reminders but it was unsure if there was face-to-face/verbal contact between researchers and participants [49]. For semi-guided interventions, six trials involved sending standardized emails periodically to encourage participants to complete the intervention [54,56], or to remind participants about the principles learned in the computer-based intervention [43-45,55]. Chiauzzi [53] sent reminder emails only if participants were not accessing the intervention for the required duration. Two trials featured weekly telephone or email-based support from a "program coach" [46] or from the researchers [55] to help participants complete the intervention or to prompt skills practice. Six trials [43-45,47,50,55] were carried out at a study site where a researcher was present to provide support and aid participants' familiarity with the intervention. One intervention involved peer interaction via online forum [56]. Three offline computer-delivered interventions involved a single session of participant-computer interaction, supplemented with hard copies of the presented material [43,44] or worksheets to complete after experiencing a stressful encounter [45]. The additional computer-delivered intervention was accessed weekly over 6 weeks and was supplemented with hard copies and a practice version of the intervention on a USB flash drive for off-site personal access [55].

Participant Characteristics

A total of 1795 participants consented and were randomized to a trial arm. Sample sizes ranged from 38 [50] to 240 [53]. Four trials had samples of ≥150 participants [45,49,51,53]. Overall, 1480 were explicitly included in analyses. Seven studies explicitly stated analysis was conducted on participants who completed pre-post intervention measures [28,42,45,48,55-57], while eight studies conducted intention-to-treat [ITT] analyses [44,46,47,49,50,52-54]. ITT was conducted through using maximum likelihood estimation [44,46], mixed-models repeated measures [49], mixed-models analysis [53], and by carrying last observation forward [52,54]. One reported separate ITT, completers, and compliers analyses [49]. Uncertainty about types of analysis was present in two publications [43,51]; 12 publications provided information regarding participant dropouts/withdrawals: dropout rates ranged from 7.2% [28] to 44.2% [54]. Five provided some reason for withdrawal; this included not receiving response to researcher's contact [44], personal time constraints [42,48,52], personal reasons [42], concerns about intervention efficacy [52], participants felt better after receiving some intervention modules [52], and participant requested face-to-face therapy instead [49].

The 10 studies describing their sample's age range included participants ranging from 17 to 51 years. In 15 trials, participants' mean age ranged from 18.37 to 28.2 years; their mean age from these was 22.6 years. All studies recruited males and females, with females being the majority in 15 studies. Gender balance varied from 50% [55] to 88.46% [54] of the sample being female. A total of 10 trials were conducted on undergraduate populations [28,42-45,47,48,50,51,53], five on both undergraduates and postgraduates [46,49,52,54,57], and two on postgraduates only [55,56]. Psychology students were overrepresented in the undergraduate studies with seven recruiting psychology undergraduates only [28,42-45,48,50] and another recruiting psychology and health sciences students [47]. Likewise, seven trials reported use of course or financial credit for participation [42-45,47,50,51,55]. The majority of trials (n=7) were conducted in HEIs in the United States [43-45,51,53,56], with three trials in Canada [28,42,46] and Australia [47,48,50], two in the United Kingdom [54,57], one in Spain [52] and Norway [49]. Further, 13 trials were conducted within one HEI [28,42-45,47,48,50,51,54-57]; the others recruited students at two [49,52], three [46], and six [53] HEIs.

Multimedia Use and Interactivity of Interventions

Limited information was provided regarding multimedia and interactivity. Text was presented in all interventions, with the use of images/graphics also reported [43,44,47-49,53,56]. Animation, music, and audio voiceovers were used in the examination anxiety intervention [57], and the social anxiety intervention utilized streaming of online videos to expose participants to an anxiety-inducing situation [52]. MoodGym [47,49,50] included interactive activities and an online workbook. Recently published studies appeared to provide more information on the presentation and interactivity of intervention content. Day [46] reported each module was presented using a range of videos, audio, pictures, and interactive activities. Mindfulness was taught through text and videos, and participants were able to choose to listen to either a male- or female-delivered 10-minute audio of meditation instruction [54]. SMART-OP [55] incorporated animation, videos, and text to create a tailored user experience, as well as using game-like interactive tasks.

Outcome Measures Used

A small number of established valid and reliable measures were used to primarily measure depression, anxiety, psychological

distress, and stress outcomes (see Table 1). Stress is an important psychological well-being outcome given that students are faced with several stressors during their studies and elevated stress can increase the risk of developing mental health difficulties [58]. All trials administered self-report measures to participants, either through hard copy or through online administration. One study administered the Trier Social Stress Test and measured associated physiological stress responses [55].

Table 1. Outcome measures used for assessment of depression, anxiety, psychological distress, and stress in the included studies.

Author	Anxiety					Depression			Psychological distress		Stress	
	ASI[a]	BAI[b]	DASS-21[c]	SAD[d]	TAI[e]	BDI[f]	CES-D[g]	DASS-21	K10[h]	PHQ-4[i]	PSS[j]	DASS-21
Arpin-Cribbie 2012	✓	✓					✓				✓	
Botella 2010				✓		✓						
Braithwaite 2007		✓				✓						
Braithwaite 2009		✓				✓						
Cavanagh 2013										✓	✓	
Chiauzzi 2008											✓	
Cukrowicz 2007		✓					✓					
Day 2013	✓		✓					✓				✓
Ellis 2011			✓					✓	✓			✓[k]
Kanekar 2010									✓[l]			
Kenardy 2003	✓						✓					
Lintvedt 2011							✓					
Orbach 2007					✓							
Radhu 2012		✓					✓				✓	
Rose 2013											✓	
Sethi 2010			✓					✓	✓			✓[k]
Taitz 2011						✓						

[a]ASI: Anxiety Sensitivity Inventory

[b]BAI: Beck Anxiety Inventory

[c]DASS-21: Depression Anxiety and Stress Scale – 21 item version

[d]SAD: Social Avoidance and Distress scale

[e]TAI: Test Anxiety Inventory

[f]BDI: Beck Depression Inventory

[g]CES-D: Center for Epidemiologic Studies Depression Scale

[h]K10: Kessler Distress Scale – 10 item version

[i]PHQ-4: Patient Health Questionnaire – 4 item version

[j]PSS: Perceived Stress Scale

[k]Data from stress subscale of DASS-21 was not reported in the published article.

[l]Shorter version of scale used to analyze data collected on K10.

Questionnaire Response Burden

Response burden reflects the amount of strain put on an individual to complete measures; factors influencing burden include length and intensity of measures and concentration required to complete them [59]. Response burden is a factor to consider in trials as participants typically complete a battery of measures at baseline and post-intervention, and potentially at more time-points during trials. Too many questions may increase burden and result in greater attrition or lower response rates [59]. We calculated the number of questions participants completed by reviewing the measures within included publications and totaling the approximated number of items within administered measures. It was estimated the measurement battery ranged from 25 [46] to 225 questions [42]. The estimated median number of questions presented to participants was 75 items.

Participant Satisfaction/Evaluation With Intervention

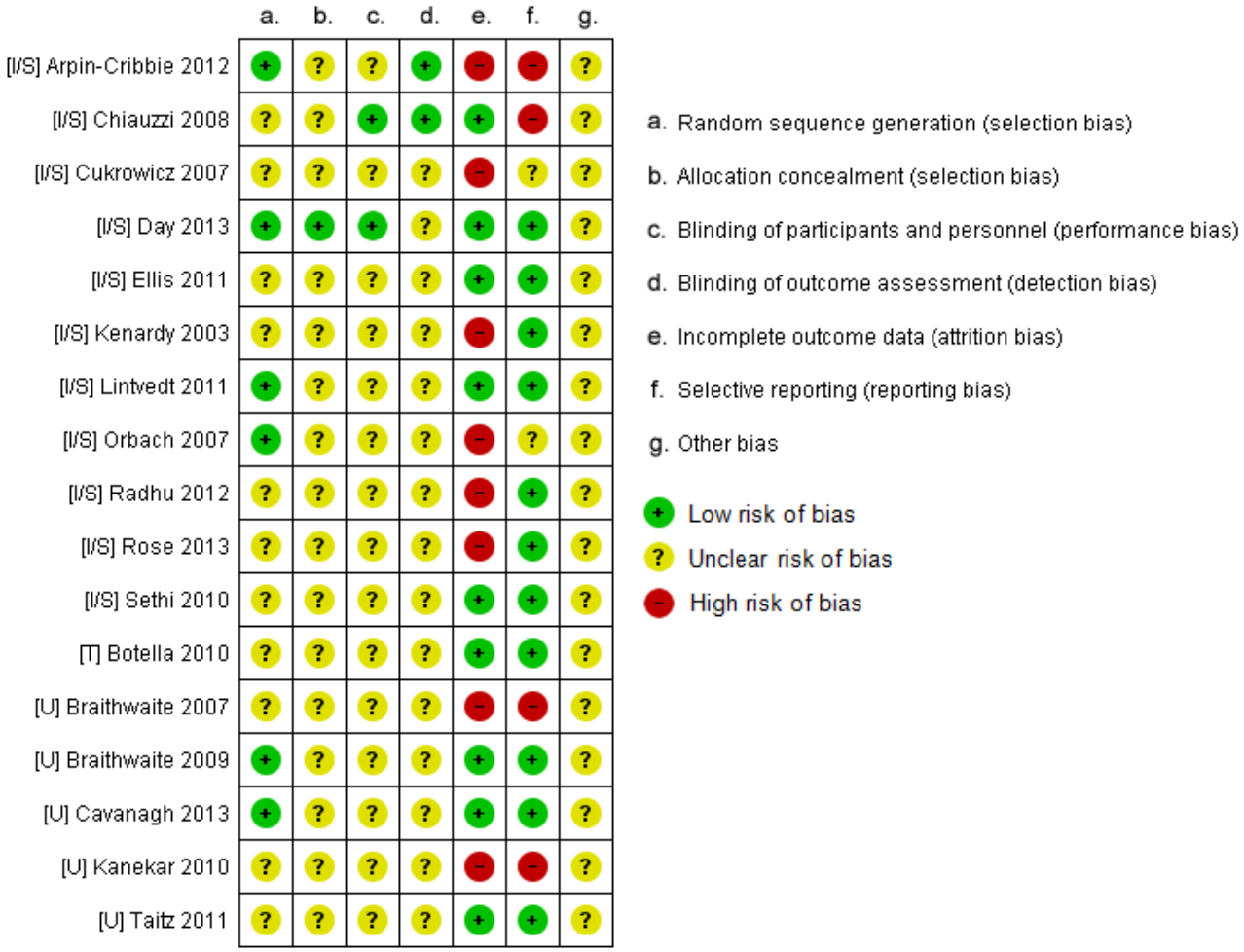
Eight studies administered a form of participant evaluation [46-49,53-55,57]. Included interventions were reported to be highly useable [55], satisfactory [53], credible [48], and to be

moderately-to-highly useful and helpful [46,47,49,54,57]. Cavanagh [54] directly asked participants if they felt the intervention had been beneficial; the majority felt the mindfulness intervention had at least some personal benefit. Day's intervention [46] underwent usability, efficiency, and acceptability testing by university students prior to being trialed [60].

Risk of Bias in Included Studies

We believed the risk of bias in included studies to be moderate—this was mostly due to publications being unclear or providing insufficient details (see Figure 2). All participants were randomized but only six studies [28,43,46,49,54,57] described their randomization method: a random number table [28], a computer-generated randomization sequence [43,46,49,54], and through tossing a coin [57]. Two studies [43,56] did not explicitly state how many participants were in each condition. It is viable to blind those collecting and/or assessing outcome data, as blinding participants can be difficult given the type of controls [14,31]. One study stated single-blindedness of participants and provided post-intervention evaluation of researchers' non-blindedness [53]; another reported single-blindedness of researcher collecting data [28]. Seven studies [28,42,45,48,55-57] explicitly did completers' analyses—overall, 208 participants were not included in analysis. Outcome data from three studies could not be extracted due to not reporting participant numbers in each condition [43,56], not reporting SD/standard error data [43,53], and assessing outcome data using a shortened version of the measure [56]. Gender balance is an issue as the majority of trials had more female than male participants. Baseline symptomology is also a potential source of bias for the review, as it may have caused some difficulties comparing intervention effectiveness in improving mental health outcomes. Trials varied in the level of mental health-related symptomology they targeted at baseline; some only recruited participants with minimal symptoms, while others wanted those experiencing elevated symptoms.

Figure 2. Breakdown of each type of risk of bias identified in the included studies.



Distribution of the Reported Data

Six studies explicitly stated their data had been checked for violations of assumptions of normality [28,45,49,53,55,57]. Two studies transformed skewed data for analysis to approximate a normal distribution [53,55], while Orbach [57] used non-parametric tests for skewed data. None of the included studies appeared to provide alternative measures of central tendency. Overall, there were 10 studies that reported skewed post-intervention on at least one primary outcome measure of interest [28,42,44-47,49,51,52,54].

Meta-Analysis for Anxiety, Depression, and Psychological Distress Outcomes

Outcome data relating to the mental health symptomology measures was not extracted from three studies due to insufficient data reporting [43,53,56]. Orbach's trial [57] was excluded from meta-analysis for anxiety outcomes, as test anxiety is considered an "extreme" reaction to examinations and is distinct from commonly diagnosable anxiety disorders [57]. Data regarding participant attrition could be extracted from two of these studies [53,57]. All mental health outcomes were continuous and scale-based, and were extracted as endpoint average scores with lower scores indicating fewer symptoms. Within the presented analyses, negative SMD values support the intervention condition.

Three analyses exploring intervention compared to inactive control, intervention compared to active control, and intervention compared to comparison intervention were conducted and are reported separately. For each type of comparison, outcomes relating to depression, anxiety, psychological distress, and stress are separately reported. For each outcome within each comparison, analyses are presented twofold: non-skewed data were analyzed first, with a secondary sensitivity analysis conducted to analyze skewed and non-skewed data on each outcome. If skewed data were present in one trial arm but not in the other, it was included in sensitivity analysis. Findings within forest plots were subgrouped by the separate measures used to measure each outcome in addition to calculation of an overall pooled effect. On all presented forest plots (see Figures), the bracketed letter before author name indicates their type: [U] universal intervention, [I/S] indicated or selective intervention, and [T] treatment intervention.

Web-Based or Computer-Delivered Intervention Compared to Inactive Control

Seven trials used this trial arm comparison to investigate effects of intervention upon anxiety outcomes. All trials were based on CBT and include four separate trials of two interventions [28,42,47,50]. Two trials reported non-skewed data—for these there was no difference between intervention and control for anxiety (n=93, 2 RCTs, pooled SMD −0.67, CI −1.59 to −0.25, Z=1.43, P=.15; I^2=66%, P=.09). Sensitivity analysis incorporated an additional five studies reporting skewed data. This analysis significantly favored the intervention (n=374, 7 RCTs, pooled SMD −0.56, CI −0.77 to −0.35, Z=5.19, P<.001; I^2=0%, P=.63; see Figure 3).

Nine trials that compared intervention to inactive control reported depression outcomes. Eight trials had CBT-based interventions and included five separate trials of two interventions [28,42,47,49,50]. Three trials reported non-skewed outcome data and significantly favored intervention (n=144, 3 RCTs, pooled SMD −0.67, CI −1.15 to −0.20, Z=2.77, P=.006; I^2=43%, P=.17). A separate sensitivity analysis included an additional six studies reporting skewed data, with this analysis significantly favoring intervention (n=712, 9 RCTs, pooled SMD −0.43, CI −0.63 to −0.22, Z=4.06, P<.001; I^2=39%, P=.11; see Figure 4).

Two trials measured psychological distress, of which one reported skewed data [54]. Cochrane Collaboration guidelines suggest forest plots should not be produced for outcomes with single studies [61]; therefore, findings from the single non-skewed trial are presented in Multimedia Appendix 5. A sensitivity analysis was performed to include the additional study, which found no difference between intervention and control (n=123, 2 RCTs, SMD −1.39, 95% CI −3.79 to 1.02, Z=1.13, P=.26). Significantly high heterogeneity was present (I^2=92%, P<.001).

Three RCTs included an outcome measure of stress. For the two studies reporting non-skewed data, there was significant favorability for intervention (n=151, 2 RCTs, pooled SMD −0.44, CI −0.77 to −0.12, Z=2.68, P=.007; I^2=0%, P=.49). A separate sensitivity analysis included the additional skewed data, which significantly favored intervention (n=217, 3 RCTs, pooled SMD −0.73, CI −1.27 to −0.19, Z=2.64, P=.008). A significant high level of heterogeneity was present (I^2=72%, P=.03).

Looking at attrition rates, participants were significantly more likely to leave the study early if they were randomly assigned to receive intervention (n=999, 11 RCTs, OR 2.73, CI 1.56-4.76, Z=3.54, P<.001; I^2=30%, P=.20; Figure 5). A total of 118 (22.7%) left the intervention condition early, compared to 52 (10.8%) in the inactive control condition.

Figure 3. Sensitivity analysis of post-intervention anxiety outcomes for intervention compared to inactive controls.

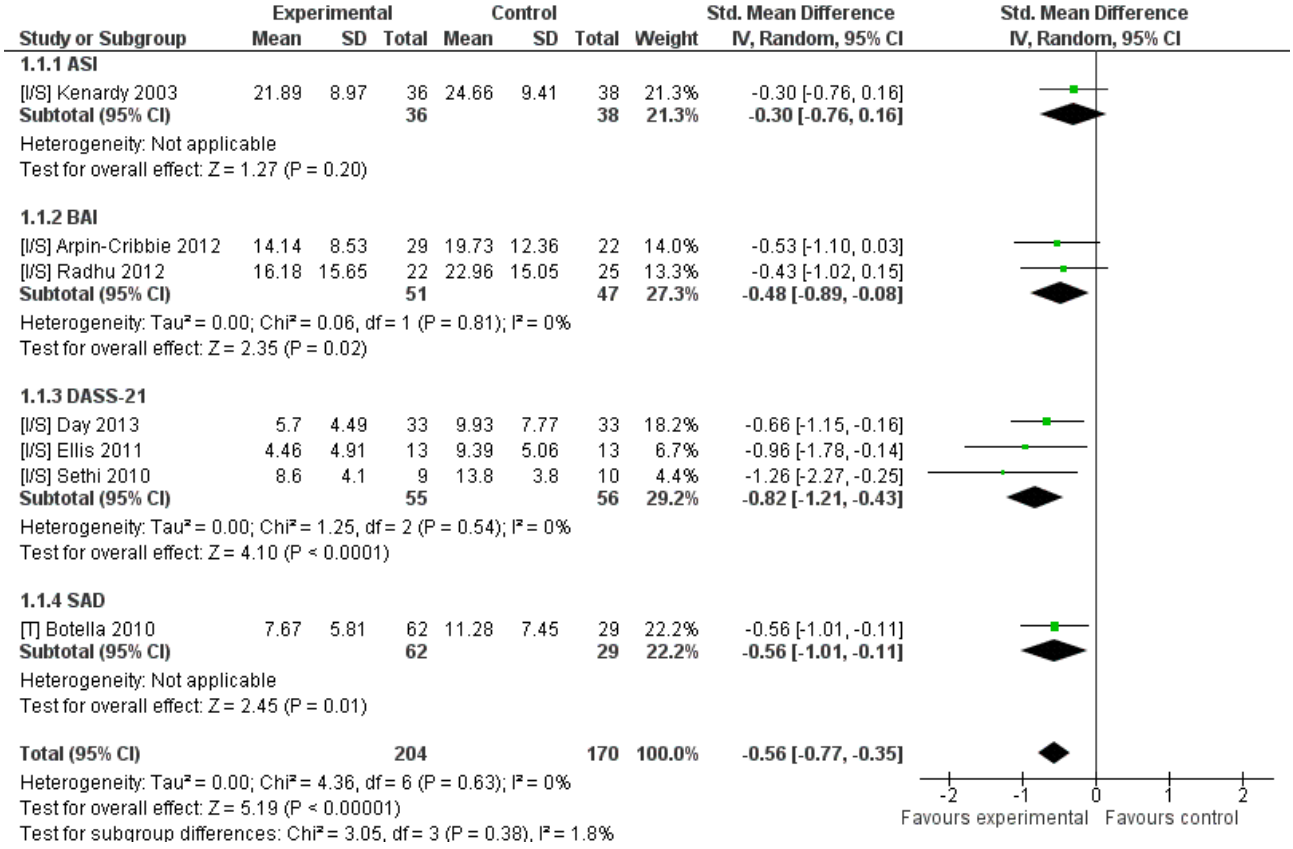


Figure 4. Sensitivity analysis of post-intervention depression outcomes for intervention compared to inactive controls.

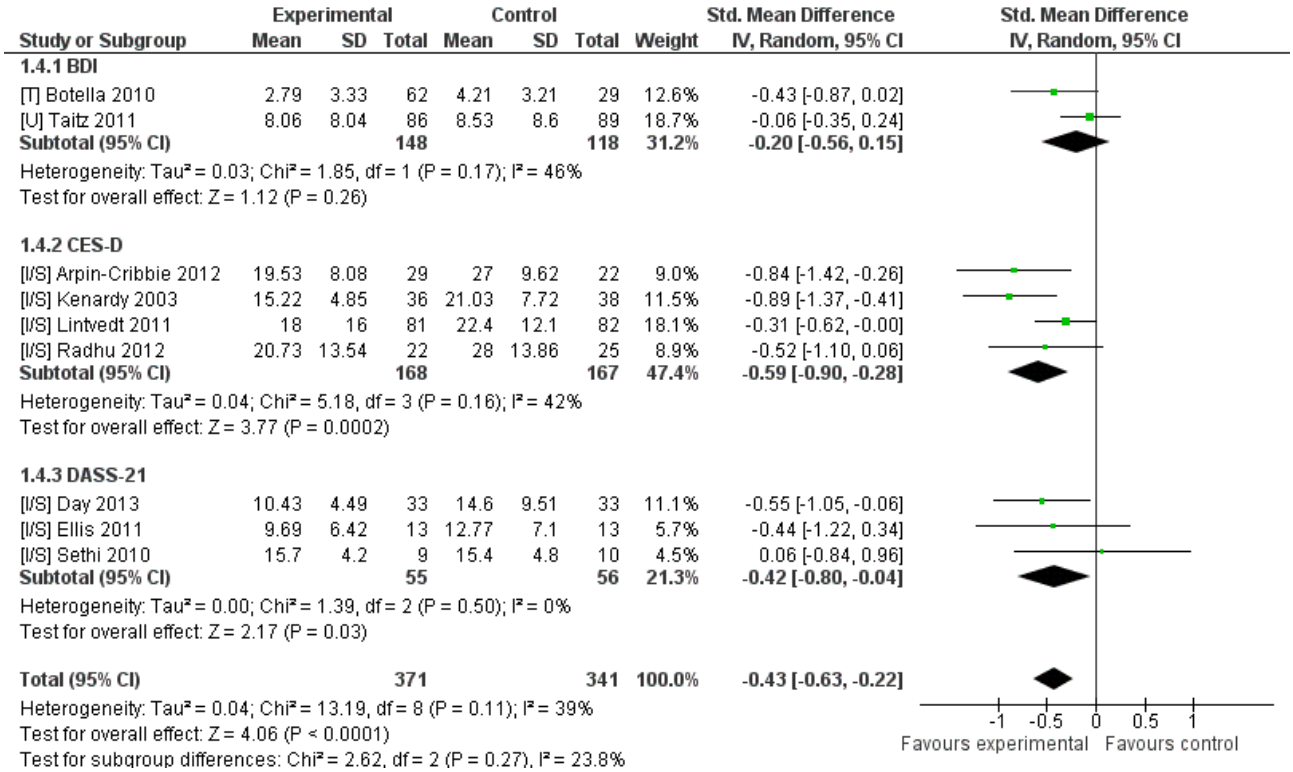
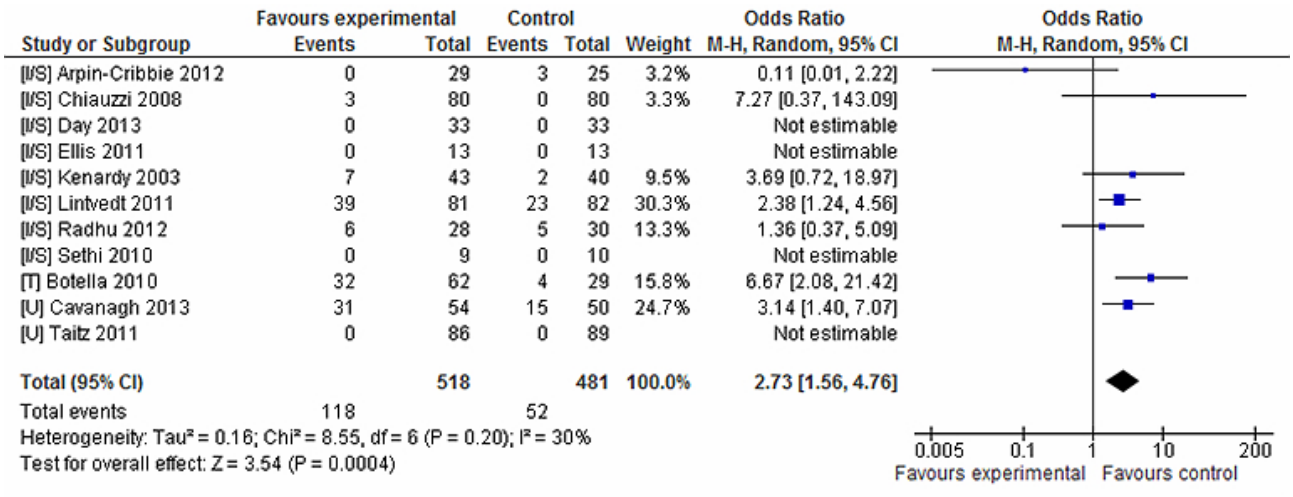


Figure 5. Attrition rates for intervention vs inactive control conditions.



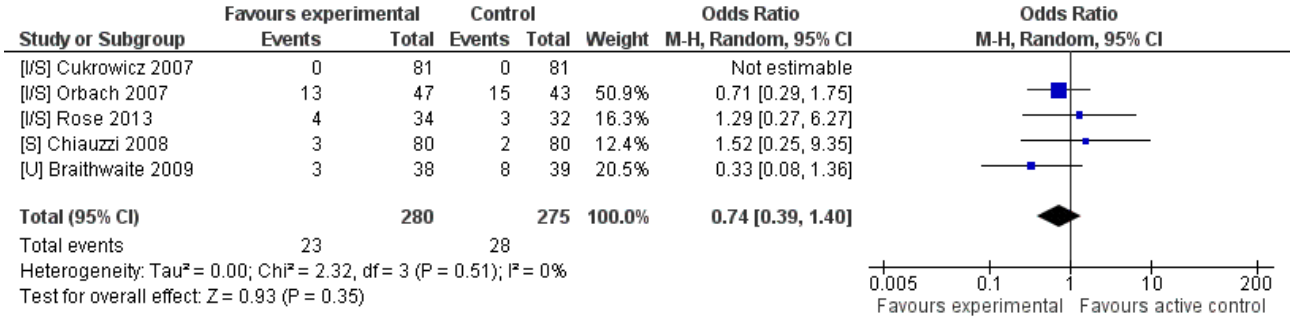
Web-Based or Computer-Delivered Intervention Compared to Active Control

There were seven trials that explicitly included an active control, but only three reported their outcome data relating to mental health outcomes of interest, or could not be included for reasons previously described. Data relating to attrition could be extracted from five of these trials. Two used the same active control in which participants viewed computer-based materials that provided descriptive information about depression and anxiety [44,45].

Two trials compared intervention to active control in investigating anxiety outcomes, both of which reported skewed data. Sensitivity analysis did not favor either intervention or active control (n=229, 2 RCTs, pooled SMD −0.18, CI −0.98 to 0.62, Z=0.45, P=.66). A high level of heterogeneity was reported (I^2=88%, P<.001). The same two trials also reported depression outcomes [44,45], which again were skewed. Sensitivity analysis did not support either condition (n=229, 2 RCTs, pooled SMD −0.28, CI −0.75 to 0.20, Z=1.14, P=.25; I^2=67%, P=.08).

Only one trial assessed psychological distress within the intervention vs active control comparison [55]. It was not subject to analysis due to being the sole study (see Multimedia Appendix 5). There were no significant differences reported between attrition in the two arms (n=555, 5 RCTs, OR 0.74, CI 0.39-1.40, Z=0.93, P=.35; I^2=0%, P=.51; see Figure 6). A total of 23 (8.2%) participants left the intervention condition early, compared to 28 (10.1%) in the active controls.

Figure 6. Attrition rates for intervention vs. active control conditions.



Web-Based or Computer-Delivered Intervention Compared to Comparison Intervention

Five trials compared the intervention to a comparison intervention. Comparison interventions were a Web-based stress management intervention [28], a face-to-face version of the intervention [52], another computer-based CBT program [43], and an online support group [47]. Sethi's trial [50] compared intervention to two comparison interventions consisting of face-to-face CBT and this combined with MoodGym. The face-to-face CBT was selected for this analysis to avoid double-counting of the intervention condition's data. Outcome data from one trial could not be extracted for analysis [43], resulting in four trials, which all reported depression and anxiety outcomes, and included two trials of MoodGym [47,50]. Sensitivity analyses were conducted for both outcomes as only one trial in each outcome reported non-skewed data (see Multimedia Appendix 5). For anxiety, neither intervention nor comparison were favored over each other (n=198, 4 RCTs, pooled SMD −0.10, CI −0.39 to 0.18, Z=0.71, P=.48; I^2=0%, P=.90; see Figure 7). Likewise for depression outcomes neither condition was favored (n=198, 4 RCTs, pooled SMD 0.33, CI −0.43 to 1.09, Z=0.85, P=.40) (see Figure 8). There was a significant high level of heterogeneity reported for depression (I^2=82%, P=.001). Only one study reported outcomes relating to psychological distress (reported in Multimedia Appendix 5). There were no differences between conditions in leaving the

study early (n=194, 4 RCTs, OR 1.18, CI 0.02-60.23, Z=0.08, P=.93; I^2=0%, P=.51). All attrition from the main intervention condition came from one study [52], wherein 32 participants left the study early. Seven (8.6%) in the comparison intervention condition left the study early.

Figure 7. Sensitivity analysis of post-intervention anxiety outcomes for intervention compared to comparison intervention.

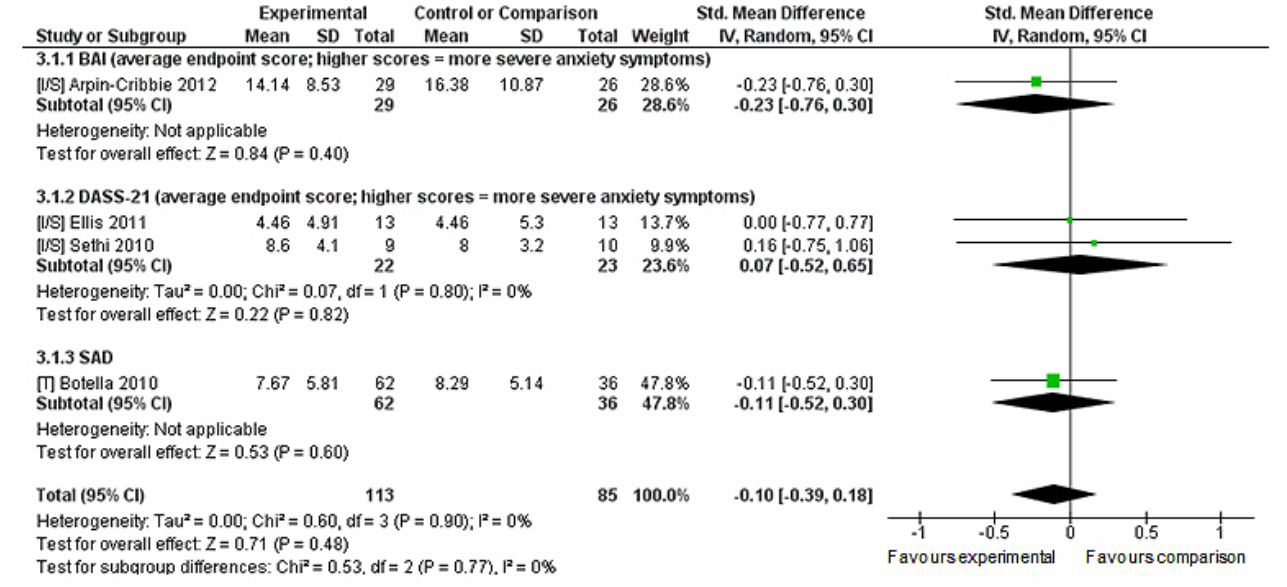
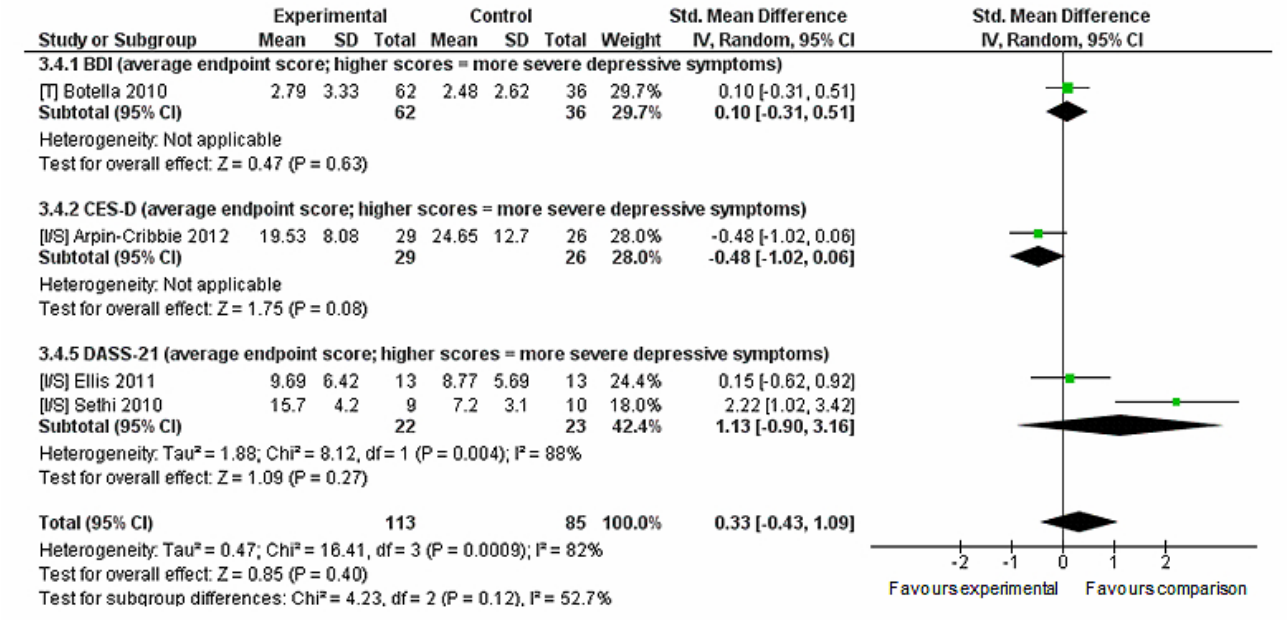


Figure 8. Sensitivity analysis of post-intervention depression outcomes for intervention compared to comparison intervention.



Additional Analyses

Given some of the methodological issues identified in the review, some additional sensitivity meta-analyses were performed. More recent publications appeared to report greater levels of methodological detail, possibly due to the research field being more established. The CONSORT-EHEALTH statement is a checklist providing a minimum list of recommendations for reporting RCTs of Internet and mobile-based interventions; it expands upon the previously published CONSORT statement [31,62]. The publication of the CONSORT-EHEALTH checklist was used as a benchmark for comparing 'older' (published ≤2011) to 'newer' (≥2012) publications. Included studies within the meta-analysis were separated based on their year of publication. These analyses could only be done for anxiety and depression outcomes in the intervention vs inactive control and vs comparison intervention comparisons due to low numbers of included trials and no differences in the publication dates in other comparisons and outcomes.

For depression outcomes in intervention compared to inactive control, a larger effect size was reported for more recent publications (n=164, 3 RCTs, pooled SMD −0.63, CI −0.94 to −0.31, Z=3.91, P<.001; I^2=0%, P=.70), than for older publications (n=548, 6 RCTs, pooled SMD −0.35, CI −0.60 to −0.09, Z=2.64, P=.008; I^2=47%, P=.09). For anxiety outcomes in the same comparison, there was little variation in the effect sizes and statistical significance in older (n=210, 4 RCTs, pooled SMD −0.60, CI −0.95 to −0.25, Z=3.37, P<.001; I^2=25%, P=.26) and newer publications (n=164, 3 RCTs, pooled SMD −0.55, CI −0.87 to −0.24, Z=3.46, P<.001; I^2=0%, P=.84). For depression and anxiety outcomes for intervention in contrast to a comparison intervention, there was only one post-2012 publication; analysis of all studies in this outcome is reported in the previous section. Looking at ≤2011 studies only, there was no difference reported between intervention and comparison for depression (n=143, 3 RCTs, pooled SMD 0.68, CI −0.33 to 1.69, Z=1.31, P=.19; I^2=82%, P=.004) or anxiety (n=143, 3 RCTs, pooled SMD −0.05, CI −0.39 to 0.28, Z=0.30, P=.76 (I^2=0%, P=.086).

Additional sensitivity analyses were also conducted to focus on trials that rewarded course credits for participation. This was not performed for the intervention vs active control comparison as all studies within this rewarded credit. Looking at studies that gave credit in the intervention vs inactive control comparison, the intervention was supported in improving anxiety outcomes (n=92, 3 RCTs, pooled SMD −0.75, CI −1.23 to −0.28, Z=3.10, P=.002; I^2=15%, P=.31) but not for depression (n=267, 4 RCTs, pooled SMD −0.16, CI −0.41 to 0.08, Z=1.33, P=.18; I^2=0%, P=.44). For studies that did not reward credit, intervention still supported anxiety (n=282, 4 RCTs, pooled SMD −0.51, CI −0.75 to −0.26, Z=4.07, P<.001; I^2=0%, P=.75) and also supported depression (n=282, 5 RCTs, pooled SMD −0.55, CI −0.78 to −0.32, Z=4.66, P<.001; I^2=26%, P=.25).

For participants who received credit in the intervention vs comparison intervention contrasts, neither condition was supported for anxiety (n=45, 2 RCTs, pooled SMD 0.07, CI −0.52 to 0.65, Z=0.22, P=.82; I^2=0%, P=.80) or depression (n=45, 2 RCTs, pooled SMD 1.13, CI −0.90 to 3.16, Z=1.09, P=.27; I^2=88%, P=.004). The same findings were repeated for participants that did not receive credit, upon anxiety (n=153, 2 RCTs, pooled SMD −0.15, CI −0.48 to 0.17, Z=0.93, P=.35; I^2=0%, P=.73) and depression outcomes (n=153, 2 RCTs, pooled SMD −0.16, CI −0.73 to 0.40, Z=0.57, P=.57; I^2=65%, P=.09).

Discussion

Principal Findings

A total of 17 studies were retrieved for this review, of which 14 were entered into meta-analysis. The majority of studies administered measures of both depression and anxiety (9/17, 53%), with two also measuring stress or psychological distress. Two studies reported targeting depression alone, with the six remaining studies reporting a mixture of outcomes. The majority were Web-based trials (n=13) with four delivered via an offline computer-delivered program. The review findings suggest Web-based and computer-delivered interventions can produce beneficial mental health outcomes in university students, supporting previous reviews of Internet and computerized interventions for depression and anxiety [14,16,40]. Our search found several recent publications not reviewed previously [4], which demonstrates the fast pace of publications in this field.

Findings demonstrated a difference in outcome data depending on the type of analyses conducted. Non-skewed data alone did not favor intervention in improving anxiety, but sensitivity analysis favored intervention when compared to inactive control. However, improvements in anxiety outcomes were not supported when intervention was compared to active control or comparison intervention. Similar findings were reported for depression outcomes. Non-skewed data for intervention compared to inactive control revealed a larger effect size (SMD −0.67) than the sensitivity analysis (SMD −0.43), suggesting skewed data can potentially affect the overall power of interventions. For psychological distress, the data did not support the intervention. The small number of studies, the different measures used, and the type of intervention complicates interpretation of findings. For stress, compared to inactive control, both meta-analyses supported intervention, with a larger effect found for sensitivity (SMD −0.73) than non-skewed analysis (−0.44). Similarly, the heterogeneity went from 0% for non-skewed analysis to 70% for sensitivity analysis, so this difference could be due to the skewed data.

When compared to inactive control, interventions appeared to be supported in improving outcomes apart from psychological distress. When compared to active control and comparison interventions, computer-delivered and Web-based interventions were not significantly supported in improving depression or anxiety. This was anticipated given that participants were still actively doing something, compared to an inactive control [40]. Neither intervention nor comparison intervention were significantly favored in meta-analysis, which may suggest some equivalency in their effect upon improving anxiety and depression outcomes. A reason this finding may have occurred could be due to the type of comparison intervention used. Two comparison interventions were face-to-face CBT, which is representative of the kind of help university students would typically receive for common mental health problems. Further research comparing these technology-based interventions to treatment-as-usual conditions would be beneficial in exploring the viability of self-guided Internet-based interventions for university students, and whether they have equivalency in comparison to the therapies young people would usually receive. Larger effect sizes within intervention vs inactive control comparisons than intervention vs active control have been reported previously in CCBT reviews [16,40]. Both active controls were identical in their content; the lack of significant effect found in the meta-analysis suggests neither intervention nor active control were more advantageous in improving outcomes. This finding may question what is the minimum level of active control needed to produce positive change.

Moderate to high heterogeneity were reported for two of the analyses comparing intervention against active control and comparison intervention. This could be due to the type of

comparison intervention or that differences in outcome data at baseline affected post-intervention symptom improvement. Grist and Cavanagh [16] identified type of control condition as being a significant moderating factor explaining heterogeneity within meta-analyses. In trials of CCBT, active controls often share some commonalities with the experimental intervention; effect sizes reported previously suggest CCBT can offer some additional small benefits in improving psychological outcomes [16]. A total of 13 studies involved CBT-based interventions, which supports findings from previous CCBT reviews [14-17]. While this continues to provide strong support for CCBT, research should explore what other evidence-based psychological and psychotherapeutic theories can be adapted to this medium [40]. It is difficult to determine which elements of the intervention produced the most beneficial effects, and there are many factors to consider, such as level of support, intervention length, the number and content of modules, and actual participant engagement.

Separating older and newer studies did appear to have an effect upon the effect sizes for depression outcomes in intervention vs inactive control comparisons, with a larger effect size found for more recent publications. Within the same comparison, there was little difference in effect sizes for anxiety, and separating the studies did not appear to add any additional insight into intervention vs comparison intervention analyses. These contrasting findings may suggest research into Internet interventions has somewhat strengthened over the years and become more methodologically sound. However, these links are tenuous given the small numbers of included trials within the separate analyses.

Future trials within university student populations should consider the effect of participant incentives and rewards upon outcomes; given that students are typically financially strained, outcomes in trials may differ from their real-world non-trial use of interventions. Separate sensitivity analyses were conducted to explore whether receiving participatory reward affected outcomes. Within the intervention vs inactive control comparison for anxiety, a larger effect size was reported for studies that did reward credit (SMD −0.75) than for those that did not (SMD −0.51). However, for depression the analysis did not support intervention in studies that rewarded credit, whereas those that did not use incentives reported a significant favoring for intervention (SMD −0.55). Sensitivity analyses for rewards within the intervention vs comparison intervention contrast reported similar findings in line with the main meta-analysis. The contrasting findings for this comparison do not allow us to precisely conclude that rewarding participants does increase an intervention's efficacy, but incentives and rewards are a factor to consider when disseminating trial findings. A meta-analysis of Web-based surveys found that incentives for participation increased individuals' motivation to start and complete the survey [63]. Similarly, college students who participated in an incentive-based online intervention for weight loss reported that financial rewards acted as a strong external motivator to lose weight and achieve weekly goals, although they also commented that the financial incentive did not influence their intrinsic motivation to participate [64]. The majority of studies that utilized participatory reward did so through providing course credit. This may differ somewhat from financial incentives but nonetheless requires consideration as students may place similar personal value upon monetary and course credit rewards. Some publications insufficiently reported their outcome data. Authors should aim to provide a CONSORT-EHEALTH statement to help report their interventions [31] so the design and content of interventions can be viewed clearly. Authors in more recent publications appeared to report more aspects of this checklist in their respective publications.

Participant dropouts were reported in 12 studies; attrition is common in these types of intervention trials [65,66]. Two studies [50,54] had similar rationales for delivering their interventions over a short timeframe—shorter interventions are associated with increased engagement and retention of participants. Baseline symptoms have been associated with attrition rates—lower depressive symptoms were positively associated with increased adherence to interventions in one review [65]. As some of the included interventions recruited participants with minimal/mild symptomology, this is an issue to consider. Only two trials assessed whether participants' levels of adherence affected their level of post-intervention improvement upon mental health, in which no associations were found [48,53]. With Internet-delivered interventions, it can be difficult to assess participants' levels of intervention engagement and there may be variation in how participant engagement is defined [31].

Participant attrition was more likely to occur in intervention groups when compared to inactive control, with no association found for comparisons to active control or comparison intervention. This was found in a review of CCBT [16]. Grist suggests the finding of no attrition differences in intervention and active control groups indicates that attrition is common in any active condition, whether it be the experimental intervention or an active control, and is not just a consequence of receiving CCBT. It may suggest some level of support is required to help participants adhere to the intervention. Only a few trials provided detail about participants' reasons for dropping out. Attrition has commonly been used as a proxy measure of participant evaluation and attitudes towards CCBT [20,48]. Interventions that do not sufficiently engage or appeal to the user may be more susceptible to dropout [48]. Interventions could potentially show positive effects due to the unengaged participants withdrawing from the study [57]; attrition may partially account for this review's positive findings. Seeking participants' reasons for disengaging from intervention is important in helping identify factors affecting adherence.

Aside from Botella's trial, which aimed to treat diagnosable social phobia [52], none of the studies explored post-intervention diagnosis of mental disorders. This is important as these interventions are used as mental health prevention and longitudinal follow-up would allow us to explore the interventions' preventative effects. Help-seeking intentions and/or behaviors were not assessed through standardized measures in any study; these interventions can subsequently affect participants' help-seeking [40]. Over a third of participants in one trial stated that as a result of the intervention they had changed their behavior, which included seeking out more information, trying self-help techniques described in the

intervention, and supporting others [49]. It is understandable that follow-up may be difficult in university students given the transient nature of university life—students may change address or leave higher education between post-intervention and follow-up periods. The timing of conducting trials is important given the fluctuating demands occurring during the academic year. Only three studies reported when post-intervention measures were administered; two of these were during examination periods [44,49] and so improvements may also be demonstrated during periods of high stress.

Just over half the interventions were semi-guided. Most of these incorporated a strategy to maintain engagement and thereby encourage adherence, such as using standardized reminders, receiving the intervention at a study site, or support from a non-therapeutic individual. We did not analyze whether there were any differences in effects between semi-guided and self-administered interventions, and cannot make assumptions about the impact of human interaction upon intervention effectiveness. A previous review found larger effect sizes for self/un-guided interventions than ones involving guidance [16].

Two interventions [46,55] had a large amount of human contact with participants. In both trials, participants received weekly contact from researchers or from program coaches to support them in completing the intervention. This kind of support provides reduced training costs compared to interventions that involve support from health care professionals, and as the program coaches were students themselves, participants may have found them relatable. Administration of trials in researcher-monitored settings could have affected participants' engagement with the intervention [14]. Johannson and Andersson [34] found increased human therapeutic support given to users was significantly associated with larger intervention effects. There was limited evaluation regarding participants' perceptions about the beneficial or therapeutic effects of human support, but nonetheless the amount of contact participants had with another person could affect intervention effectiveness.

Mental health outcomes were assessed using a small number of well-established continuous measures aligned with diagnostic criteria. This made comparisons in the meta-analysis less complicated; however, having several measures can increase statistical heterogeneity [67]. We attempted to counteract this by investigating intervention effects by subgrouping each type of measure within each outcome, and looking separately at the overall pooled effect. By doing this, we could explore measurement comparisons for each outcome, which did show some variation in the different measures used for the same outcomes.

The overwhelming presence of skewed data in the included studies affected the quality of the available evidence. Skewed data has been reported previously in a review of computer-delivered interventions for reducing alcohol consumption [68]. Almost all included studies reported the mean and standard deviation from outcome measures, and none reported alternative measures of central tendency. Only a minority had transformed skewed data or used non-parametric tests. The meta-analyses reported a vast quantity of heterogeneity, which hinders their generalizability, and the differences in the scoring range of measures may be a reason why it occurred. For example, the two psychological distress measures varied on their scoring range: the PHQ-4 (Patient Health Questionnaire) was a brief measure where scores range from 0 to 12, while scores on the K10 (Kessler Distress Scale) range from 0 to 40. Large heterogeneity has been reported previously in reviews of Internet-delivered and computer-based interventions for depression [40,69]. Richards and Richardson [69] suggest eligibility criteria can be a cause of heterogeneity. This is possible given the variation in the baseline symptomology eligibility criteria of included participants. Some trials recruited participants experiencing minimal to moderate levels of depression, anxiety, or stress [45-47,50]; within some of the same analyses, there were participants who were included if they were experiencing elevated symptoms [48,49]. This variation in symptomology may affect the overall power of the included interventions.

Small sample sizes were apparent. The smallest sample involved 38 participants, within which there were four arms, of which two contained nine participants each [50]. There was limited detail about power calculations to recruit appropriate sample sizes. The forest plots show studies with smaller samples were associated with larger confidence intervals and are less reliable than larger samples. Coupling this with the considerable skew means the findings need to be approached with caution. Completers analysis may bias the calculated effectiveness of interventions as these analyses are likely to produce larger outcome effects [70]. ITT analysis helps avoid selection bias that can occur if only those completing measures at all study time-points are analyzed [71].

The use of participation reminders requires consideration. Interventions trialed in the included studies may not have reminders when administered in a non-trial context. Three studies trialed MoodGym, a freely available online resource that any member of the public can sign up to. In this context, general public users do not receive reminders to complete the intervention—unlike in two included studies [47,50] where participants completed it in a monitored setting.

Funnel plots were briefly inspected to explore possible presence of publication bias; these did not appear to show any unusual asymmetry. This was approached with caution as funnel plot asymmetry should ideally be used when ≥10 studies are in analysis [72]. The majority of studies reported positive outcomes on at least one relevant mental symptomology measure. We did not include non-peer reviewed studies and so did not include unpublished data. As reported previously by Farrer [4], not all may have been designed for university students—instead they were sampled to opportunistically trial out the intervention and they may have some differences to the ideal target population. Participants in some studies were already experiencing minimal symptoms upon enrolment, meaning it is problematic to determine how much of an effect the intervention had upon reducing developmental risk of ill mental health. For example, intervention participants in one trial [44] reported a mean pre-post intervention decline of <3 points on the BDI (Beck Depression Inventory); at baseline, participants were already classified as having minimal depressive symptoms. It is difficult

to address the significance of this small decrease in already minimal symptomology, and the preventative effect of interventions is further complicated by limited follow-up. No studies assessed utilization of mental health services or diagnosis of mental disorders as an outcome measure, making it difficult to know if interventions reduced the risk of developing a mental disorder or affected mental health service use. For the meta-analyses, only post-intervention short-term data were used due to limited long-term follow-up. We are unsure about the long-term maintenance of improvements in outcomes.

Participants in seven studies received course or financial credit for participation [42-45,47,50,51,55] and eight samples were recruited from psychology degree courses. In sensitivity analyses, one comparison for depression (interventions vs inactive control) did not support the intervention, whereas it did in the overall analysis. This may bias findings as those who participated for credit are likely be different from students who seek help without an reward incentive for doing so. Likewise psychology students may be more knowledgeable about mental health and the trial process, and thus more receptive to interventions. However, the effects may be greater in students who were not aware of the possibilities of CBT/evidence-based approaches to improve mood. The overrepresentation of psychology students may account for the gender imbalance in recruitment [73]. Young male adults are frequently cited as being less likely to seek out help for their mental health [74,75], and it has been suggested Internet-based interventions could reach out to men [75]. Researchers need to reach out to students in other disciplines and also recruit more males to their trials. Another factor to consider relates to the age range of participants. Unlike Farrer [4], we did not have age as inclusion criteria for the review. The average age calculated from 15 included studies was 22.6 years, and some samples included older adults. This deviates from the traditional age range of university students, and older students may have different mental health needs than typically aged students. Given this, the findings may not be fully generalizable to younger students. Future research would benefit by focusing on sampling students within the 18-25 year age range typical of student populations, or consider age as a moderating factor of intervention effectiveness within this population.

A moderate risk of bias was calculated mostly due to insufficient details reported about trial methodology and outcome measures, meaning we were unclear about several risk of bias outcomes. Only a minority of studies reported their randomization method; this has been reported previously in reviews of CCBT, technology-based interventions, and interventions to improve help-seeking and stigmatizing attitudes and beliefs in university students [4,16,23,76]. Grading the blindness of participants in included studies may be irrelevant given the nature of the types of intervention and trial design [40]. Some studies insufficiently reported their data, which affects the quality of the available evidence. Reporting methodological factors, such as randomization method, concealment, and the blinding of research personnel, is essential to judging trial quality. Researchers in this field are becoming more aware of using CONSORT-EHEALTH guidelines in their publications [31], which addresses several of these methodological factors.

While all included studies explored the statistical significance of outcome data, only a few looked into whether improvements were clinically significant. The few that calculated these found intervention participants showed a higher level of reliable and clinically significant improvement compared to controls [28,46,52,57]. Calculating this provides additional value about the recovery status of participants. It would also be useful to explore whether the improvements reported in the outcome measures correspond to participants' perceptions, as there has been disagreement between severity of symptoms reported on a common depression measure and participants' actual verbal description of symptom severity [77]. This could be done by asking them whether they felt the intervention helped their mental well-being, and might help to address the apparent overreliance on focusing on psychometric measures. One qualitative study found students felt use of an online resource helped them manage their mental well-being during periods of psychological distress [78].

Implications for Practice

As the intervention vs comparison intervention analyses suggested some level of equivalence in outcomes, individuals working in student health, such as welfare advisors and counsellors, may be considering online and technology-based resources they can use to support their students. Some universities do appear interested in using online resources, as several British HEIs have incorporated Web-based interventions into their welfare services, such as the "CALM/Relief" series [79]. None of the included studies assessed whether these interventions had outcomes upon students' academic performance. This is likely to be an important outcome for policymakers given the reputation of their institutions. The best improvements in mental health outcomes may be achieved by combining self-help with face-to-face support [19].To help address the increased demand for university-based counselling, online resources could be used as a support tool by university students while waiting to see a relevant professional [78]. Similarly, these resources could also be used as an adjunct by students in between counselling appointments.

Implications for Research

Future research needs to consider sufficient sample sizes required for trials, and address the skewed data present in outcome data by either transforming it or using alternative tests. Measurements of help-seeking intentions and behavior, as well as aspects of mental health literacy, would be highly useful in future research as online interventions are often promoted as an alternative to seeking face-to-face help or preventing onset of ill mental health [23]. Researchers would benefit from collaborating with the student population to understand what measurable outcomes are important to them; as these young people are in higher education to obtain a qualification, it is expected that academic performance and retention would be salient outcomes. Mental health difficulties can significantly impair students' academic performance and social functioning; future research should incorporate outcomes reflecting these domains. Gaining user evaluation of interventions through qualitative methods such as interviews and focus groups would also be highly useful in attaining feedback to address the worth

of the intervention and to make interventions more appropriate for student needs [20].

Limitations

All studies were coded by one author (EBD) and were discussed as necessary with CG. The use of one coder may have unintentionally biased the results. There is the possibility that relevant publications may have been missed in the search. However, the search was conducted on several databases and updated through a repeat search, as it had taken some time to conduct the review. Likewise, Farrer's review [4] was searched for additional publications. For meta-analysis, we could not extract data from three included trials, meaning the pool of data from included interventions was smaller. Similarly for the anxiety meta-analyses, measures that may reflect certain distinct aspects of anxiety disorders, such as anxiety sensitivity and social anxiety, were incorporated into one analysis for all anxiety outcomes, which may also have induced bias. The studies trialing the same three interventions had slight variation in how they individually conducted and how participants accessed the intervention. Lintvedt [49] coupled MoodGym with an information-only website, meaning participants received additional information not delivered in the other MoodGym trials [47,50]. The type of intervention may have influenced the reported heterogeneity. In their meta-analysis of Internet-delivered CBT for depression and anxiety, Spek [14] found higher heterogeneity in treatment interventions compared to ones focused on prevention. For our review, there was only one intervention that could clearly be defined as treatment; however, there was variation in the type of universal and selective/indicated interventions being trialed. The level of human support and contact within included interventions is another aspect affecting participant-intervention engagement, which may have impacted effect sizes [14].

Trials of mobile apps for improving mental health outcomes were not included in this review, as it was felt these were still an emerging technology at the time. University students may be a group likely to use apps as they also present many of the same benefits as computer-based/Web-based interventions, but could be more accessible given the popularity of smartphones and tablets. Farrer's review [4] was explored for app-based interventions. A recent review of mental health apps for smartphones/tablets found only five apps that had been trialed [80], one of which was trialed on a student population [81]. However, as found with several in the present review, this trial's methodology and data were not reported clearly and it is unclear whether the intervention was a smartphone app.

Several studies analyzed conducted completers analyses, which may bias review findings as these analyses are likely to produce larger outcome effects [70]. All interventions used different content and multimedia, which could affect how much participants interacted with the intervention and subsequently their effectiveness [23]. It is difficult to know whether improvements produced by both intervention and active control conditions would have been maintained in the long-term due to limited follow-up. Given that some active controls/comparison interventions produced similar outcome effects to the intervention being trialed, consideration is needed regarding the minimum intervention needed to produce effective change in outcomes. Use of active controls may result in difficulty in understanding the true effect of the experimental intervention upon outcomes [70].

Interventions from different theoretical approaches were combined together for the meta-analysis. Limited numbers of non-CBT trials meant separate analyses exploring different approaches could not be conducted. Although there were only a small number of non-CBT trials within meta-analyses, this could potentially skew findings and so future reviews may want to separately analyze outcomes based on the theoretical underpinning of interventions. Random Effects Models were used for all analyses; however, this may induce bias as it places larger significance on smaller studies [82]. Many trials involved small samples, meaning this bias may have occurred. Finally, no-treatment control and wait-list controls were collapsed into one comparison category (inactive control) for analysis. There were seven trials using wait-list and four using a no-treatment control. This could affect findings as those assigned to wait-list control would have been expecting to receive intervention at some point and may show improvements in their symptomology due to expectation effects.

Conclusions

Overall, this review provides some cautious findings that suggest online and computer-delivered interventions can potentially be beneficial in improving depression, anxiety, and psychological distress outcomes in university students. These interventions are not a panacea for all, but do provide an easily implemented health promotion and prevention strategy that can be easily reached by university students. The benefits of these interventions may potentially help HEIs in promoting good mental health and well-being to its population and support students' academic performance [83]. However, trials in this review did not assess students' academic performance before or after receiving intervention. The findings support the effectiveness of the adaptation of CBT into self-guided, Internet-delivered interventions. However, several methodological shortcomings, including small sample sizes and a large amount of skewed data, mean the findings need to be treated with a high degree of caution. As concluded in a meta-analysis of psycho-educational mental health interventions [70], there needs to be more investigation into the factors influencing intervention effectiveness. Further participant feedback is encouraged to evaluate online and computer-based interventions and to help further tailor interventions to university student populations.

Acknowledgments

EBD is supported by funding from Mental Health Research UK, who were not involved in this review. RM is partly supported by NIHR CLAHRC (National Institute for Health Research Collaboration for Leadership in Applied Health Research and Care)

East Midlands and NIHR Mindtech HTC (Health Technology Co-operative). The authors would like to thank Professor Clive Adams and Hannah Jones (The University of Nottingham) for their help in the review process and analysis, and Dr Lou Farrer (Australian National University) for providing advice regarding classifying the types of intervention.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search terms used in online databases.

[PDF File (Adobe PDF File), 305KB - jmir_v16i5e130_app1.pdf]

Multimedia Appendix 2

List of studies excluded from the review.

[PDF File (Adobe PDF File), 211KB - jmir_v16i5e130_app2.pdf]

Multimedia Appendix 3

PRISMA checklist.

[PDF File (Adobe PDF File), 249KB - jmir_v16i5e130_app3.pdf]

Multimedia Appendix 4

Summary of Web-based and computer-delivered interventions to improve depression, anxiety, psychological distress, and stress conducted in university student populations included in the present review.

[PDF File (Adobe PDF File), 119KB - jmir_v16i5e130_app4.pdf]

Multimedia Appendix 5

Non-skewed data that could not be incorporated into meta-analyses due to being sole study for specific outcomes of interest.

[PDF File (Adobe PDF File), 297KB - jmir_v16i5e130_app5.pdf]

References

1. Zivin K, Eisenberg D, Gollust SE, Golberstein E. Persistence of mental health problems and needs in a college student population. J Affect Disord 2009 Oct;117(3):180-185. [doi: 10.1016/j.jad.2009.01.001] [Medline: 19178949]
2. Ibrahim AK, Kelly SJ, Adams CE, Glazebrook C. A systematic review of studies of depression prevalence in university students. J Psychiatr Res 2013 Mar;47(3):391-400. [doi: 10.1016/j.jpsychires.2012.11.015] [Medline: 23260171]
3. Macaskill A. The mental health of university students in the United Kingdom. British Journal of Guidance & Counselling 2013 Aug;41(4):426-441. [doi: 10.1080/03069885.2012.743110]
4. Farrer L, Gulliver A, Chan JK, Batterham PJ, Reynolds J, Calear A, et al. Technology-based interventions for mental health in tertiary students: systematic review. J Med Internet Res 2013;15(5):e101 [FREE Full text] [doi: 10.2196/jmir.2639] [Medline: 23711740]
5. Andrews B, Wilding JM. The relation of depression and anxiety to life-stress and achievement in students. Br J Psychol 2004 Nov;95(Pt 4):509-521. [doi: 10.1348/0007126042369802] [Medline: 15527535]
6. Hysenbegasi A, Hass SL, Rowland CR. The impact of depression on the academic productivity of university students. J Ment Health Policy Econ 2005 Sep;8(3):145-151. [Medline: 16278502]
7. Russell G, Shaw S. A study to investigate the prevalence of social anxiety in a sample of higher education students in the United Kingdom. J Ment Health 2009 Jan;18(3):198-206. [doi: 10.1080/09638230802522494]
8. Eisenberg D, Golberstein E, Gollust SE. Help-seeking and access to mental health care in a university student population. Med Care 2007 Jul;45(7):594-601. [doi: 10.1097/MLR.0b013e31803bb4c1] [Medline: 17571007]
9. Gulliver A, Griffiths KM, Christensen H. Perceived barriers and facilitators to mental health help-seeking in young people: a systematic review. BMC Psychiatry 2010;10:113 [FREE Full text] [doi: 10.1186/1471-244X-10-113] [Medline: 21192795]
10. 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 Jun;40(2):82-97. [doi: 10.1080/16506073.2011.573807]
11. Carey KB, Scott-Sheldon LA, Elliott JC, Bolles JR, Carey MP. Computer-delivered interventions to reduce college student drinking: a meta-analysis. Addiction 2009 Nov;104(11):1807-1819 [FREE Full text] [doi: 10.1111/j.1360-0443.2009.02691.x] [Medline: 19744139]

12. Ryan ML, Shochet IM, Stallman HM. Universal online interventions might engage psychologically distressed university students who are unlikely to seek formal help. Advances in Mental Health 2010 Aug;9(1):73-83. [doi: 10.5172/jamh.9.1.73]
13. Taylor CB, Luce KH. Computer- and internet-based psychotherapy interventions. Current Directions in Psychol Sci 2003 Feb;12(1):18-22. [doi: 10.1111/1467-8721.01214]
14. 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]
15. Kaltenthaler E, Parry G, Beverley C, Ferriter M. Computerised cognitive-behavioural therapy for depression: systematic review. Br J Psychiatry 2008 Sep;193(3):181-184 [FREE Full text] [doi: 10.1192/bjp.bp.106.025981] [Medline: 18757972]
16. Grist R, Cavanagh K. Computerised cognitive behavioural therapy for common mental health disorders, what works, for whom under what circumstances? A systematic review and meta-analysis. J Contemp Psychother 2013;43(4):243-251. [doi: 10.1007/s10879-013-9243-y]
17. Foroushani PS, Schneider J, Assareh N. Meta-review of the effectiveness of computerised CBT in treating depression. BMC Psychiatry 2011;11:131 [FREE Full text] [doi: 10.1186/1471-244X-11-131] [Medline: 21838902]
18. Escoffery C, Miner KR, Adame DD, Butler S, McCormick L, Mendell E. Internet use for health information among college students. J Am Coll Health 2005;53(4):183-188. [doi: 10.3200/JACH.53.4.183-188] [Medline: 15663067]
19. Rickwood D, Bradford S. The role of self-help in the treatment of mild anxiety disorders in young people: an evidence-based review. Psychol Res Behav Manag 2012;5:25-36 [FREE Full text] [doi: 10.2147/PRBM.S23357] [Medline: 22427736]
20. Mitchell N, Gordon PK. Attitudes towards computerized CBT for depression amongst a student population. Behav Cognit Psychother 2007 May 14;35(04):421. [doi: 10.1017/S1352465807003700]
21. Royal College of Psychiatrists. London: Royal College of Psychiatrists. 2011. Mental health of students in higher education URL: http://www.rcpsych.ac.uk/publications/collegereports/cr/cr166.aspx [accessed 2014-04-01] [WebCite Cache ID 6OVSdeAzn]
22. Lintvedt OK, S rensen 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 Jul 03;26(2-4):239-258. [doi: 10.1080/15228830802096705]
23. Gulliver A, Griffiths KM, Christensen H, Brewer JL. A systematic review of help-seeking interventions for depression, anxiety and general psychological distress. BMC Psychiatry 2012;12:81 [FREE Full text] [doi: 10.1186/1471-244X-12-81] [Medline: 22799879]
24. Yager Z, O'Dea JA. Prevention programs for body image and eating disorders on university campuses: a review of large, controlled interventions. Health Promot Int 2008 Jun;23(2):173-189 [FREE Full text] [doi: 10.1093/heapro/dan004] [Medline: 18263883]
25. Bewick BM, Trusler K, Barkham M, Hill AJ, Cahill J, Mulhern B. The effectiveness of Web-based interventions designed to decrease alcohol consumption--a systematic review. Prev Med 2008 Jul;47(1):17-26. [doi: 10.1016/j.ypmed.2008.01.005] [Medline: 18302970]
26. Givi H, Imani H, Agh A, Mohammadipourrik N, Mehrabadi S. Efficiency of computerized cognitive behavioral therapy versus clinical intervention for the treatment of major depression translated from Farsi. Koomesh: Journal of Semnan University of Medical Sciences 2012;13(2):218-224.
27. Kenardy J, McCafferty K, Rosa V. Internet-delivered indicated prevention for anxiety disorders: Six-month follow-up. Clinical Psychologist 2006;10(1):39-42. [doi: 10.1080/13284200500378746]
28. Arpin-Cribbie C, Irvine J, Ritvo P. Web-based cognitive-behavioral therapy for perfectionism: a randomized controlled trial. Psychother Res 2012 Mar;22(2):194-207. [doi: 10.1080/10503307.2011.637242] [Medline: 22122217]
29. Arpin-Cribbie CA, Irvine J, Ritvo P, Cribbie RA, Flett GL, Hewitt PL. Perfectionism and psychological distress: A modeling approach to understanding their therapeutic relationship. J Rat-Emo Cognitive-Behav Ther 2008 Feb 27;26(3):151-167. [doi: 10.1007/s10942-007-0065-2]
30. Higgins JPT, Deeks JJ. Selecting studies and collecting data. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions. Chichester, England: Wiley-Blackwell; 2008.
31. 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]
32. The Cochrane Collaboration. Review Manager Software (RevMan) (Version 5.1). Copenhagen: The Nordic Cochrane Centre; 2012. URL: http://tech.cochrane.org/Revman [accessed 2014-05-09] [WebCite Cache ID 6PRD63U1D]
33. Mrazek PJ, Haggerty RJ. In: Mrazek PJ, Haggerty RJ, editors. Reducing Risks for Mental Disorders: Frontiers for Preventive Intervention Research. Washington DC, USA: The National Academies Press; 1994.
34. Johansson R, Andersson G. Internet-based psychological treatments for depression. Expert Rev Neurother 2012 Jul;12(7):861-870. [doi: 10.1586/ern.12.63] [Medline: 22853793]
35. Higgins JPT, Altman DG. Assessing risk of bias in included studies. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions. Chichester, England: Wiley-Blackwell; 2008.

36. Altman DG, Bland JM. Statistics notes: detecting skewness from summary information. BMJ 1996 Nov 9;313(7066):1200 [FREE Full text] [Medline: 8916759]
37. Hedges LV, Vevea JL. Estimating effect size under publication bias: Small sample properties and robustness of a random effects selection model. Journal of Educational and Behavioral Statistics 1996 Jan 01;21(4):299-332. [doi: 10.3102/10769986021004299]
38. Deeks JJ, Higgins JPT, Altman DG. Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions. Chichester, England: Wiley-Blackwell; 2008.
39. Hofmann SG, Sawyer AT, Witt AA, Oh D. The effect of mindfulness-based therapy on anxiety and depression: A meta-analytic review. J Consult Clin Psychol 2010 Apr;78(2):169-183 [FREE Full text] [doi: 10.1037/a0018555] [Medline: 20350028]
40. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. Cogn Behav Ther 2009;38(4):196-205. [doi: 10.1080/16506070903318960] [Medline: 20183695]
41. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003 Sep 6;327(7414):557-560 [FREE Full text] [doi: 10.1136/bmj.327.7414.557] [Medline: 12958120]
42. Radhu N, Daskalakis ZJ, Arpin-Cribbie CA, Irvine J, Ritvo P. Evaluating a Web-based cognitive-behavioral therapy for maladaptive perfectionism in university students. J Am Coll Health 2012;60(5):357-366. [doi: 10.1080/07448481.2011.630703] [Medline: 22686358]
43. Braithwaite SR, Fincham FD. ePREP: Computer based prevention of relationship dysfunction, depression and anxiety. Journal of Social and Clinical Psychology 2007 May;26(5):609-622. [doi: 10.1521/jscp.2007.26.5.609]
44. Braithwaite SR, Fincham FD. A randomized clinical trial of a computer based preventive intervention: replication and extension of ePREP. J Fam Psychol 2009 Feb;23(1):32-38. [doi: 10.1037/a0014061] [Medline: 19203157]
45. Cukrowicz KC, Joiner TE. Computer-based intervention for anxious and depressive symptoms in a non-clinical population. Cogn Ther Res 2007 Apr 4;31(5):677-693. [doi: 10.1007/s10608-006-9094-x]
46. Day V, McGrath PJ, Wojtowicz M. Internet-based guided self-help for university students with anxiety, depression and stress: a randomized controlled clinical trial. Behav Res Ther 2013 Jul;51(7):344-351. [doi: 10.1016/j.brat.2013.03.003] [Medline: 23639300]
47. Ellis LA, Campbell AJ, Sethi S, O'Dea BM. Comparative randomized trial of an online cognitive-behavioral therapy program and an online support group for depression and anxiety. J Cyber Ther Rehabil 2011;4(4):461-467.
48. Kenardy J, McCafferty K, Rosa V. Internet-delivered indicated prevention for anxiety disorders: A randomized controlled trial. Behav Cognit Psychother 2003;31(3):279-289. [doi: 10.1017/S1352465803003047]
49. Lintvedt OK, Griffiths KM, Sørensen K, Østvik 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;20(1):10-27. [doi: 10.1002/cpp.770] [Medline: 21887811]
50. Sethi S, Campbell AJ, Ellis LA. The use of computerized self-help packages to treat adolescent depression and anxiety. Journal of Technology in Human Services 2010 Aug 31;28(3):144-160. [doi: 10.1080/15228835.2010.508317]
51. Taitz I. Learning lucid dreaming and its effect on depression in undergraduates. International Journal of Dream Research 2011;4(2):117-126. [doi: 10.11588/ijodr.2011.2.9123]
52. Botella C, Gallego MJ, Garcia-Palacios A, Guillen V, Baños RM, Quero S, et al. An Internet-based self-help treatment for fear of public speaking: a controlled trial. Cyberpsychol Behav Soc Netw 2010 Aug;13(4):407-421. [doi: 10.1089/cyber.2009.0224] [Medline: 20712499]
53. Chiauzzi E, Brevard J, Thum C, Thurn C, Decembrele S, Lord S. MyStudentBody-Stress: an online stress management intervention for college students. J Health Commun 2008 Sep;13(6):555-572. [doi: 10.1080/10810730802281668] [Medline: 18726812]
54. Cavanagh K, Strauss C, Cicconi F, Griffiths N, Wyper A, Jones F. A randomised controlled trial of a brief online mindfulness-based intervention. Behav Res Ther 2013 Sep;51(9):573-578. [doi: 10.1016/j.brat.2013.06.003] [Medline: 23872699]
55. Rose RD, Buckey JC, Zbozinek TD, Motivala SJ, Glenn DE, Cartreine JA, et al. A randomized controlled trial of a self-guided, multimedia, stress management and resilience training program. Behav Res Ther 2013 Feb;51(2):106-112. [doi: 10.1016/j.brat.2012.11.003] [Medline: 23262118]
56. Kanekar A, Sharma M, Atri A. Enhancing social support, hardiness, and acculturation to improve mental health among Asian Indian international students. Int Q Community Health Educ 2009;30(1):55-68. [doi: 10.2190/IQ.30.1.e] [Medline: 20353927]
57. Orbach G, Lindsay S, Grey S. A randomised placebo-controlled trial of a self-help Internet-based intervention for test anxiety. Behav Res Ther 2007 Mar;45(3):483-496. [doi: 10.1016/j.brat.2006.04.002] [Medline: 16814744]
58. Bitsika V, Sharpley CF, Rubenstein V. What stresses university students: An interview investigation of the demands of tertiary studies. Australian Journal of Guidance and Counselling 2010;20(1):41-54. [doi: 10.1375/ajgc.20.1.41]
59. Rolstad S, Adler J, Rydén A. Response burden and questionnaire length: is shorter better? A review and meta-analysis. Value Health 2011 Dec;14(8):1101-1108. [doi: 10.1016/j.jval.2011.06.003] [Medline: 22152180]

60. Currie SL, McGrath PJ, Day V. Development and usability of an online CBT program for symptoms of moderate depression, anxiety, and stress in post-secondary students. Computers in Human Behavior 2010 Nov;26(6):1419-1426. [doi: 10.1016/j.chb.2010.04.020]
61. Schunemann H, Oxman A, Higgins JPT, Vist G, Glasziou P, Guyatt G. Presenting results and 'summary of findings' tables. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions. Chichester, England: Wiley-Blackwell; 2008.
62. Schulz KF, Altman DG, Moher D. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010;c332:340. [doi: 10.1136/bmj.c332]
63. Göritz A. Incentives in Web studies: Methodological issues and a review. International Journal of Internet Science 2006;1(1):58-70.
64. Davy BM, Potter KL, Dennis Parker EA, Harden S, Hill JL, Halliday TM, et al. Feasibility, effectiveness, and perceptions of an Internet-and incentive-based behavioral weight loss intervention for overweight and obese college freshmen: A mixed methods approach. OJPM 2013;3(7):429-440. [doi: 10.4236/ojpm.2013.37058]
65. 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]
66. Waller R, Gilbody S. Barriers to the uptake of computerized cognitive behavioural therapy: a systematic review of the quantitative and qualitative evidence. Psychol Med 2009 May;39(5):705-712. [doi: 10.1017/S0033291708004224] [Medline: 18812006]
67. Lieb K, Völlm B, Rücker G, Timmer A, Stoffers JM. Pharmacotherapy for borderline personality disorder: Cochrane systematic review of randomised trials. Br J Psychiatry 2010 Jan;196(1):4-12 [FREE Full text] [doi: 10.1192/bjp.bp.108.062984] [Medline: 20044651]
68. Khadjesari Z, Murray E, Hewitt C, Hartley S, Godfrey C. Can stand-alone computer-based interventions reduce alcohol consumption? A systematic review. Addiction 2011 Feb;106(2):267-282. [doi: 10.1111/j.1360-0443.2010.03214.x] [Medline: 21083832]
69. Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. Clin Psychol Rev 2012 Jun;32(4):329-342. [doi: 10.1016/j.cpr.2012.02.004] [Medline: 22466510]
70. Donker T, Griffiths KM, Cuijpers P, Christensen H. Psychoeducation for depression, anxiety and psychological distress: a meta-analysis. BMC Med 2009;7:79 [FREE Full text] [doi: 10.1186/1741-7015-7-79] [Medline: 20015347]
71. Jüni P, Altman DG, Egger M. Systematic reviews in health care: Assessing the quality of controlled clinical trials. BMJ 2001 Jul 7;323(7303):42-46 [FREE Full text] [Medline: 11440947]
72. Sterne JAC, Egger M, Moher D. Addressing reporting biases. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions. Chichester, England: Wiley-Blackwell; 2008.
73. Willyard C. Men: A growing minority?. 2011. URL: http://www.apa.org/gradpsych/2011/01/cover-men.aspx [accessed 2013-11-21] [WebCite Cache ID 6LIglXb7Q]
74. Ridge D, Emslie C, White A. Understanding how men experience, express and cope with mental distress: where next? Sociol Health Illn 2011 Jan;33(1):145-159. [doi: 10.1111/j.1467-9566.2010.01266.x] [Medline: 21039617]
75. Ellis LA, Collin P, Hurley PJ, Davenport TA, Burns JM, Hickie IB. Young men's attitudes and behaviour in relation to mental health and technology: implications for the development of online mental health services. BMC Psychiatry 2013;13:119 [FREE Full text] [doi: 10.1186/1471-244X-13-119] [Medline: 23601273]
76. Yamaguchi S, Wu SI, Biswas M, Yate M, Aoki Y, Barley EA, et al. Effects of short-term interventions to reduce mental health-related stigma in university or college students: a systematic review. J Nerv Ment Dis 2013 Jun;201(6):490-503. [doi: 10.1097/NMD.0b013e31829480df] [Medline: 23719324]
77. Malpass A, Shaw A, Kessler D, Sharp D. Concordance between PHQ-9 scores and patients' experiences of depression: a mixed methods study. Br J Gen Pract 2010 Jun;60(575):e231-e238 [FREE Full text] [doi: 10.3399/bjgp10X502119] [Medline: 20529486]
78. Davis-McCabe C. CALM: Two years on - students' experience of the CALM approach. Association for University and College Counselling Journal 2009;November:25-27.
79. Ultrasis. Leading universities purchase Relief series and Beating the Blues programmes. 2008 Feb 13. URL: http://ultrasisplc.com/news/article/170.html [accessed 2013-11-28] [WebCite Cache ID 6LSrMPQZL]
80. Donker T, Petrie K, Proudfoot J, Clarke J, Birch MR, Christensen H. Smartphones for smarter delivery of mental health programs: a systematic review. J Med Internet Res 2013;15(11):e247 [FREE Full text] [doi: 10.2196/jmir.2791] [Medline: 24240579]
81. Grassi A, Gaggioli A, Riva G. New technologies to manage exam anxiety. Stud Health Technol Inform 2011;167:57-62. [Medline: 21685642]
82. Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ 2011;343:d4002. [Medline: 21784880]
83. Regehr C, Glancy D, Pitts A. Interventions to reduce stress in university students: a review and meta-analysis. J Affect Disord 2013 May 15;148(1):1-11. [doi: 10.1016/j.jad.2012.11.026] [Medline: 23246209]

Abbreviations

ASI: Anxiety Sensitivity Inventory
BAI: Beck Anxiety Inventory
BDI: Beck Depression Inventory
CBT: Cognitive Behavioral Therapy
CCBT: Computerized Cognitive Behavioral Therapy
CES-D: Center for Epidemiologic Studies Depression Scale
CI: confidence interval
DASS-21: Depression Anxiety and Stress Scale – 21 item version
HEI: higher education institution
ITT: intention-to-treat
K-10: Kessler Distress Scale – 10 item version
PHQ-4: Patient Health Questionnaire – 4 item version
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSS: Perceived Stress Scale
RAM: Random Effects Model
RCT: Randomized Controlled Trial
SAD: Social Avoidance and Distress scale
SMD: Standardized Mean Difference
TAI: Test Anxiety Inventory
[U]: universal intervention
[I/S]: indicated or selective intervention
[T]: treatment intervention

Edited by G Eysenbach; submitted 29.11.13; peer-reviewed by A Gulliver, C Spates; comments to author 11.02.14; revised version received 14.03.14; accepted 30.03.14; published 16.05.14.

Please cite as:
Davies EB, Morriss R, Glazebrook C
Computer-Delivered and Web-Based Interventions to Improve Depression, Anxiety, and Psychological Well-Being of University Students: A Systematic Review and Meta-Analysis
J Med Internet Res 2014;16(5):e130
URL: http://www.jmir.org/2014/5/e130/
doi:10.2196/jmir.3142
PMID:24836465

Original Paper

Effectiveness of a Web-Based Solution-Focused Brief Chat Treatment for Depressed Adolescents and Young Adults: Randomized Controlled Trial

Jeannet Kramer[1], PhD; Barbara Conijn[1], MSc; Pien Oijevaar[2], MA; Heleen Riper[3,4], PhD

[1]Trimbos Institute (Netherlands Institute of Mental Health and Addiction), Utrecht, Netherlands

[2]Developlay BV, Publishing media for children, Utrecht, Netherlands

[3]Department of Clinical Psychology and EMGO Institute, VU University, Amsterdam, Netherlands

[4]Innovation Incubator, Leuphana University, Lueneburg, Germany

Corresponding Author:
Jeannet Kramer, PhD
Trimbos Institute (Netherlands Institute of Mental Health and Addiction)
Post Office Box 725
Utrecht, 3500 AS
Netherlands
Phone: 31 (0)30 2959380
Fax: 31 (0)30 2971111
Email: jkramer@trimbos.nl

Abstract

Background: Up to 9% of young people suffer from depression. Unfortunately, many in need of help remain untreated. The Internet offers anonymous ways to help depressed youth, especially those who are reluctant to search for help because of fear of stigma.

Objective: Our goal was to evaluate the effectiveness of an individual chat treatment based on Solution-Focused Brief Therapy (SFBT) to young individuals aged 12-22 years with depressive symptoms by comparing it to a waiting list control group.

Methods: For this study, 263 young people with depressive symptoms were randomized to the Web-based SFBT intervention, PratenOnline, or to a waiting list control condition. The chat treatment was delivered by trained professionals. Groups were compared on depressive complaints as measured by the Center for Epidemiologic Studies Depression Scale (CES-D) after 9 weeks and 4.5 months. For the chat group only, changes in depressive symptoms at 7.5 months after baseline were explored.

Results: The experimental SFBT condition (n=131) showed significantly greater improvement than the waiting list condition (n=132) in depressive symptoms at 9 weeks and 4.5 months on the CES-D, with a small between group effect size at 9 weeks (d=0.18, 95% CI -0.10 to 0.47) and a large effect size at 4.5 months (d=0.79, 95% CI 0.45-1.08). The percentage of participants showing a reliable and clinically significant change in depression was significantly larger for the SFBT intervention at 4.5 months only (28.2% vs 11.4% for the waiting list, P<.001, number needed to treat=6). At 7.5 months, the SFBT group showed further improvements. However, results have to be considered carefully because of high attrition rates.

Conclusions: The Web-based SFBT chat intervention of PratenOnline was more effective than a waiting list control group in reducing depressive symptoms, and effects were larger at follow-up then at post-treatment. More studies are needed to find out if outcomes will be replicated, especially for those younger than 18 year old.

Trial Registration: Netherlands Trial Register: NTR 1696; http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=1696 (Archived by WebCite at http://www.webcitation.org/6DspeYWrJ).

(J Med Internet Res 2014;16(5):e141) doi:10.2196/jmir.3261

KEYWORDS

depression; randomized controlled trial; Internet; Solution Focused Brief Therapy; young adults

Introduction

Background

Depression is among the most common mental health problems in young people. About 5.6% of youth aged 13-18 [1], and 9% of those aged 18-24 years [2] suffer from depression. Depression early in life can have serious implications on social, educational, and family functioning and is an important predictor of suicidal behavior [3].

Despite the high prevalence of depression in youth and the possible serious implications on their lives, depression in young people is often unrecognized and undertreated [4]. Young people are not inclined to seek help for depression, and referral to treatment at mental health services is a bridge too far for most of them [5]. Perceived stigma and concern about family member responses are important barriers [6]. The reluctance of many depressed young people to engage with mental health services [7] highlights the importance of low threshold and easily accessible interventions. The Internet offers such an opportunity. The anonymity of the Internet reduces fear of stigma [8] and fits well into the "digital lifestyle" of young people.

An increasing number of Web-based services and interventions are available for children, adolescents, and young adults ranging from self-help materials to online treatments. Research on youth and young adults indicates that Web-based interventions can be effective in reducing depressive complaints [9,10]. However, outcomes of randomized controlled trials have had mixed results, with some showing better outcomes compared to a waiting list for males only [11], the whole sample [12], or compared to an active control condition [13]. Some did not show significant differences between group effects when compared to a waiting list [14] or to an active treatment control group [15]. Most studies focused on Cognitive Behavioral Therapy or Problem Solving Therapy, but none of the studies on Web-based treatments are based on Solution-Focused Brief Therapy (SFBT) [16]. SFBT shifts the focus away from problem formation and problem resolution, to participants' future goals, strengths, and resiliencies. In SFBT, a professional collaborates with the client to look for solutions to obtain goals and strongly stresses the client's autonomy and competencies to achieve them. SFBT is a widely used therapeutic approach in coaching, couples therapy, and psychotherapy. According to several meta-analyses and reviews, it has positive effects in a broad range of settings and problem areas [17-21]. In the most recent and comprehensive review, five studies focus on depression as an outcome [21]. One study focused on mildly depressed college students [22] and found that one session of SFBT was as effective as one session of interpersonal therapy with a significant decrease in depressive symptoms. Other studies on SFBT with adult populations showed that SFBT was related to a reduction of depressive symptoms over time, and comparable outcomes to short-term psychodynamic therapy [23], past-focused treatment [24], common factors therapy [25], and a treatment based on the Hazeldon model in a group of substance abusers [26]. None of these studies were about Web-based interventions.

Current Study

In this paper, we present the results of a trial on a Web-based anonymous SFBT chat intervention for depressed adolescents and young adults aged 12-22 years. The trial was started after a pilot study showed promising results: a positive evaluation by participants and a decrease from pre- to post-intervention with a large effect size (d=1.32) [27]. The trial was conducted to find out if the SFBT chat intervention was effective in reducing depressive symptoms compared to a waiting list control group. To the best of our knowledge, no randomized controlled trial has been published on the effectiveness of Web-based treatments based on SFBT, in adolescents or adults.

Methods

Study Design

A randomized controlled trial with two parallel groups was conducted to evaluate the effectiveness of the PratenOnline chat intervention (Chat) by comparing it to a waiting list control group (WL). This study was registered with the Netherlands Trial Register (NTR 1696). Ethical approval was granted by an independent medical ethics committee (Centrale Commissie Mensgebonden Onderzoek, CCMO No. NL25219.097.08).

Study Population

Participants were young people with depressive symptoms who fulfilled the following criteria: (1) 12-22 years of age, (2) had access to a computer and Internet, (3) had a CES-D score of 22 or higher (the cut-off to detect possible cases of depression among adolescents) [28], (4) gave informed consent, and (5) completed a baseline questionnaire. Applicants were excluded when there was an indication of suicidal ideation with intent and plan as measured with an item of the Quick Inventory of Depressive Symptomatology-Self rated (QIDS-SR) [29].

Recruitment

Participants for the study were recruited through articles in newspapers, and banners and links placed on relevant websites for youth and on Facebook. Young people interested in participating were referred to the PratenOnline website for information about the study and to fill in a screening questionnaire to check the criteria for involvement. Those aged 12-17 years with a CES-D score of 22 or higher were invited to fill in a Web-based informed consent form and baseline questionnaire. Candidates younger than 18 years also needed written parental consent. After inclusion, participants were automatically randomized to one of two conditions: the PratenOnline chat intervention (Chat) or the waiting list control condition (WL). Random allocation was automated by a computer program without interference of the intervention supervisor or researcher. Participants were informed by email of their allocation, and the Chat participants were asked to schedule their first chat session via the intervention website. During the study, the PratenOnline chat intervention was exclusively accessible for applicants participating in the study. Blinding of participants, therapists, and researchers was not possible due to the design of this study. During the trial, participants in both conditions were allowed to seek additional help if they wished.

The Intervention

The intervention is a brief Web-based Solution-Focused synchronous chat intervention for young people aged 12-22 with depressive symptoms called PratenOnline (Talking online) [30]. It is offered by a mental health care foundation for youth in the Netherlands (Stichting Jeugdriagg Noord Holland Zuid) and has been online since 2004. The chat consists of individual real-time chat sessions with a trained health care professional in a secured chat room. During the sessions, SFBT techniques [31] are used by the therapist, starting with asking the "miracle" question (ie, a question that asks the patient to envision and describe how the future will be different when the problem is gone), setting goals, looking for strengths or solutions, keeping the focus on what is going well or better, giving compliments, looking for exceptions to the problem, and asking the client to indicate on scales from 1-10 what progress is made in obtaining goals. At the end of each chat session, the participant decides if their intervention goal has been reached. If not, a new chat session is scheduled with the therapist. The intervention is accessible anonymously, without cost for participants and available during weekdays, (late) nights, and weekends. After registration, the participant can choose three possible dates for a chat with a therapist. The confirmation of the chat can be found after logging in to the personal mailbox on the intervention website. No reminders could be sent to an email address outside of this secured environment because of anonymity reasons. The chat intervention follows the principles of SFBT [16] and is performed according to the guidelines of the European Brief Therapy Association (EBTA; the EBTA Solution Focused Practice Definitions [32]). A chat session takes about one hour. The intention is to keep the number of chats limited to five, but more sessions are delivered when needed.

Conditions: The Waiting List

The waiting list (WL) group did not receive access to the chat intervention. They could participate after the waiting period of 4.5 months.

Assessments

Assessments took place before randomization (baseline, t0), 9 weeks (t1), and 4.5 months after baseline (t2). At 7.5 months after baseline (t3), a last follow-up measurement took place, exclusively for participants in the Chat condition, to measure effects at longer term. All assessments consisted of self-reported Web-based questionnaires and took about 15 minutes to complete. Email reminders were sent after 7 days if necessary. To stimulate response, participants received a voucher of €10 for each completed questionnaire (t1, t2, and t3).

Primary Outcome Measure: Depressive Symptoms

Symptoms of depression in the past week were assessed with the 20-item CES-D [33,34]. The total score ranges from 0-60, with higher scores reflecting more depressive symptoms. Construct validity and reliability of the CES‐D are well established for the paper-and-pencil, computerized, and Internet versions [28,35]. In our study, Cronbach alpha ranged from .75-.81.

Additional Measures

At baseline, demographic characteristics (ie, sex, age, educational level, daily activity, living situation, ethnic background), duration of the psychological complaints (ie, how long the current complaints had been present), and professional help received ever before and at present were assessed. At t1, professional help and use of medication were measured. Attendance of chats was automatically measured by client Web statistics.

Power

Originally, the trial was powered to detect clinically significant health gains expressed as a standardized effect size of a medium size (difference between groups of at least d=0.40) in a one-sided test with an alpha of .05 and a power (1-beta) of .80. The results reported in this paper, however, are based on more conservative two-tailed tests.

Analyses

All analyses were performed on the intention-to-treat sample with missing values imputed. The expectation-maximization (EM) method was used to impute missing data. It imputes values by maximum-likelihood estimation using the observed data in an iterative process [36]. T tests, chi-square tests, and non-parametric Mann-Whitney U tests (P<.05) were used to assess whether the randomization had resulted in two comparable groups at baseline and whether any differential loss to follow-up had occurred. Logistic regression was used (backward method) to find predictors of completing questionnaires and attending chats in the Chat condition (0=no chats, 1=one or more chats).

Change scores based on EM imputation were used to analyze differences between groups at 9 weeks and 4.5 months (a positive score means improvement). Variables on which conditions differed significantly at baseline were regarded as relevant confounders when causing a change of 10% in the regression coefficient for condition when added to the regression model [37]. While no relevant confounders were found, results of independent samples t tests are shown.

As attrition was rather high, sensitivity analyses were run to study the robustness of the estimates of EM imputation, using the multiple imputation Predictive Mean Matching method (PMM) in Stata (creating 100 datasets). PMM combines the standard linear regression and the nearest-neighbor imputation approaches. Predictors of outcome and missingness were taken into account to impute missing CES-D outcomes. Analyses were performed in a multiple imputation framework. Also data of completers of questionnaires were analyzed.

Magnitudes of intervention effects were estimated using Cohen's d [38]. Within group effect sizes were first calculated for each condition separately ($(M_{t0}–M_{t1}) / SD_{t0}$) and subsequently the between group effect size delta d by subtracting the effect size of the WL group from that of the Chat group. For Cohen's d, an effect size of 0.2 to 0.3 may be regarded as a small effect, around 0.5 as a medium effect, and 0.8 to infinity as a large effect.

The proportion of participants showing reliable and clinically significant improvement [39] was defined by an improvement of 5 points in combination with a score lower than 22 on the CES-D (cut-off based on Cuijpers et al, 2008) [28]. Differences between groups were tested with chi-square tests. The number needed to treat was calculated as 1/success rate difference [40].

The change from baseline to 7.5 months (t3) was explored in the Chat condition only by means of a one sample two-sided *t* test, comparing the change in CES-D from baseline to 7.5 months.

The analyses were performed using SPSS (version 19.0) and Stata (version 11.1).

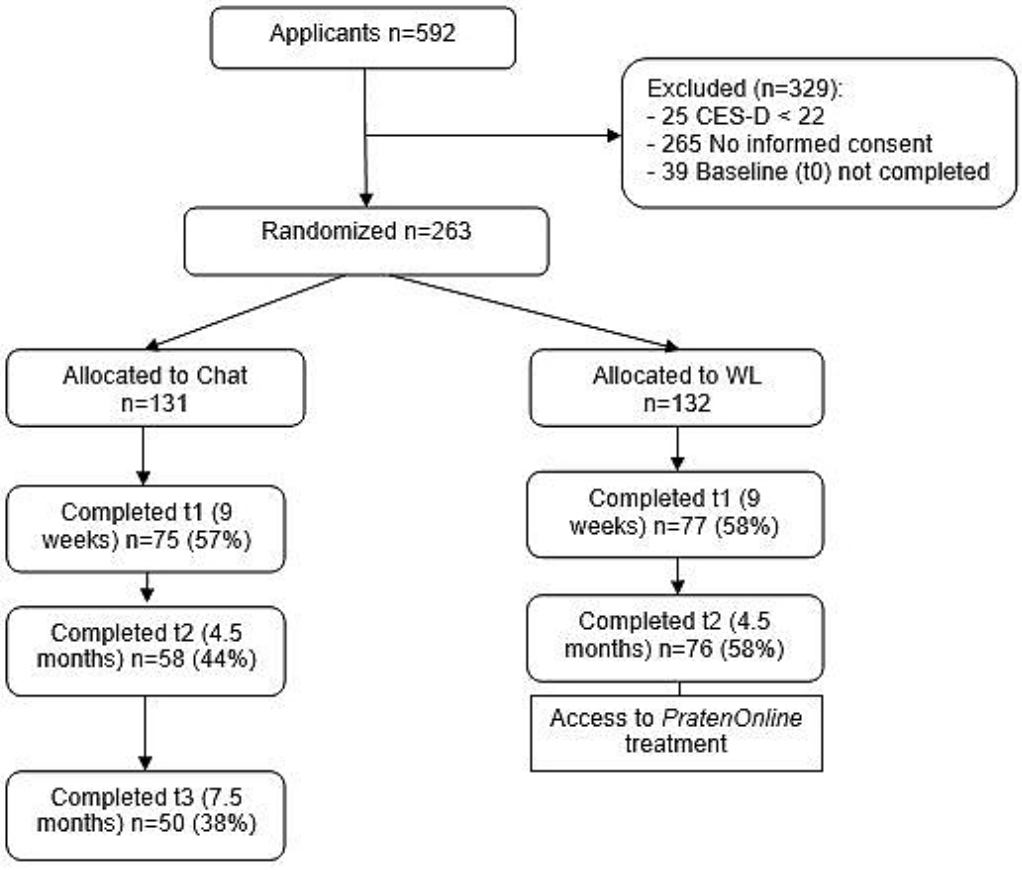
Results

Participants

Participants were recruited from August 9, 2009, until January 24, 2010. Most participants were recruited via Internet (520/592, 87.8%). Others applied on advice of a person (61/592, 10.3%) or after reading about it in a magazine or newspaper (11/592, 1.9%). Figure 1 shows the flow of participants through the trial.

Of the 592 young people who applied, 263 (44.4%) were included in the study. Reasons for non-inclusion were lack of informed consent (265/329, 80.5%), not completing the t0 questionnaire (39/329, 11.9%), and a CES-D depression score lower than 22 (25/329, 7.6%). Only 10 (3.8%) participants included were between 12 and 17 years of age (five assigned to each arm). Of the 253 applicants between 12 and 17 years of age, 243 were excluded, either because they did not return their parents' consent (227/243, 93.4%) or had a CES-D score lower than 22 (16/243, 6.6%).

Figure 1. Flow of participants.



Demographic Characteristics Participants

Baseline demographic, psychosocial, and clinical characteristics are shown in Table 1. Most participants were female (207/263, 78.7%). Over two thirds was still in school or studying (187/263, 71.1%). There were no differences between groups at baseline. Also there were no differences between groups in professional help received at t1 (χ^2_1=0.30, *P*=.59) and t2 (χ^2_1=0.07, *P*=.79) or in the use of antidepressant medication at t1 (χ^2_1=0.15, *P*=.70) and t2 (χ^2_1=0.02, *P*=.89).

Table 1. Baseline characteristics (n=263).

Characteristics	Chat n=131	WL n=132	All N=263	Statistics
Female, n (%)	104 (79.4)	103 (78.0)	207 (78.7)	χ^2_1=0.07, *P*=.79
Age, mean (SD)	19.4 (1.6)	19.6 (1.8)	19.5 (1.7)	t_{261}=0.48, *P*=.48
Age groups, years, n (%)				χ^2_1=0.00, *P*=.99
12-17	5 (3.8)	5 (3.8)	10 (3.8)	
18-22	126 (96.2)	127(96.2)	253 (96.2)	
Education level [a] , n (%)				χ^2_2=2.75, *P*=.25
Low	75 (57.3)	78 (59.1)	153 (58.2)	
Middle	50 (38.2)	42 (31.8)	92 (35.0)	
High	6 (4.6)	12 (9.1)	18 (6.8)	
Daily activity, n (%)				χ^2_2=0.57, *P*=.75
Student (high school)	91 (69.5)	96 (72.7)	187 (71.1)	
Paid job	20 (15.3)	20 (15.2)	40 (15.2)	
Other	20 (15.3)	16 (12.1)	36 (13.7)	
Living situation, n (%)				χ^2_3=0.66, *P*=.88
With parents	82 (62.6)	78 (59.1)	160 (60.8)	
With partner	18 (13.7)	17 (12.9)	35 (13.3)	
Alone	15 (11.5)	18 (13.6)	33 (12.5)	
With others	16 (12.2)	19 (14.4)	35 (13.3)	
Had professional help before, n (%)	68 (51.9)	67 (50.8)	135 (51.3)	χ^2_1=0.04, *P*=.85
Had professional help at baseline, n (%)	26 (19.8)	18 (13.6)	44 (16.7)	χ^2_1=1.82. *P*=.18
Ethnic background[b], n (%)	19 (14.5)	12 (9.1)	31 (11.8)	χ^2_1=1.85, *P*=.17
Duration psychological complaints in years, n (%)	0.70 (1.1)	0.52 (0.8)	0.61 (1.0)	t_{261}=-1.57, *P*=.12
CES-D depression score, mean (SD)	39.5 (8.6)	39.7 (7.1)	39.6 (7.9)	$t_{251.8}$=0.26, *P*=.79

[a]Education: lower=primary education or lower general secondary education, middle=intermediate vocational or high school, high=higher vocational education or university.

[b]Non-western immigrants when one or both parents is born in Africa, Latin America, or Asia (including Turkey and excluding Indonesia, Japan, and Dutch East Indies).

Attrition

A total of 42.2% (111/263) of the participants did not complete t1, and 49.0% (129/263) did not complete t2. The groups did not differ at t1 in returning completed questionnaires (χ^2_1=0.03, *P*=.86). Some statistically significant differences at baseline were detected between participants who completed measurements and those who did not. At t1, non-completers were more often males (χ^2_1=6.51, *P*=.01), lived with their parents more often (χ^2_1=7.17, *P*=.007), had a longer history of mood problems (χ^2_1=4.85, *P*=.03). At t2, non-completers were more often in the Chat group (χ^2_1=4.65, *P*=.03), were less often at school or studying (χ^2_1=4.42, *P*=.04), and had at baseline more thoughts about suicide (χ^2_1=6.01, *P*=.01). These results indicate that loss to follow-up was not completely at random. Assessment at 7.5 months (t3) was not completed by 61.8% (81/131) in the Chat group. At t3, non-completers differed from completers in depressive complaints at baseline: non-completers more often had a score of 40 or higher on the CES-D (χ^2_1=5.35, *P*=.02).

Effect of the Intervention: Primary Outcome, Depressive Symptoms

The results for the CES-D outcomes for the intention-to-treat sample are depicted in Table 2, and mean CES-D scores per measurement are shown in Figure 2. The results of *t* tests show that depressive symptoms decreased significantly more in the Chat condition from baseline to 9 weeks with a small between group effect size (*d*=0.18, 95% CI -0.10 to 0.47) and from baseline to 4.5 months with a large between group effect size of *d*=0.79 (95% CI 0.45-1.08). The sensitivity analyses with PMM imputed data showed significant differences between

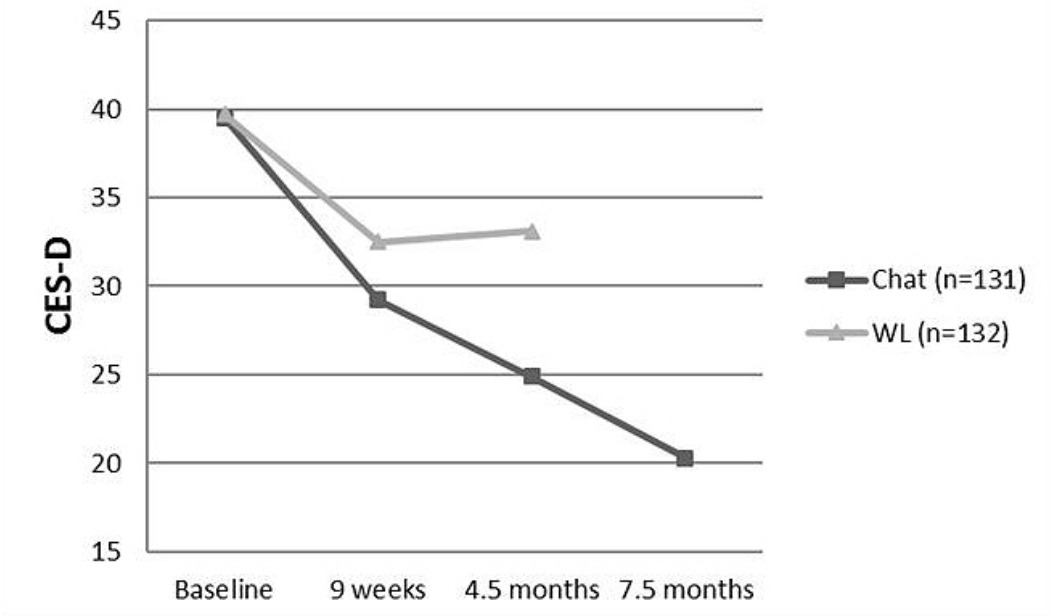
groups only at 4.5 months, with effect sizes a bit lower than the EM outcomes, but of the same magnitude being again small at 9 weeks and large at 4.5 months. Results of the analyses including only completers of questionnaires show significant differences between groups at 4.5 months, with again a large effect size in favor of the Chat condition. No significant differences were found at 9 weeks.

Table 2. Means and estimates for depression score (CES-D) at 9 weeks (t1) and 4.5 months (t2) follow-up: intention-to-treat (EM imputation) and completers only (CO) analysis.

	Chat			WL			Between group test of change scores			
	N	Mean (SD)	*d*	N	Mean (SD)	*d*	Δd[a]	*t*	df	*P* value, 2-sided
T0	131	39.49 (8.58)		132	39.74 (7.13)					
T1 EM	131	29.20 (10.66)	1.20	132	32.51 (9.68)	1.01	0.18	2.70	261	.007
T2 EM	131	24.86 (8.51)	1.72	132	33.09 (9.69)	0.93	0.79	6.39	261	<.001
T1 PMM	131	29.78 (14.41)	1.13	132	32.78 (12.95)	0.98	0.16	1.78	151.8	.08
T2 PMM	131	26.36 (14.81)	1.53	132	32.99 (13.26)	0.95	0.58	3.45	134.9	.001
T1 CO	74	29.49 (12.11)	1.07	77	33.00 (10.95)	0.94	0.14	1.53	149	.13
T2 CO	58	24.66 (10.87)	1.57	76	33.37 (11.32)	0.82	0.75	3.66	132	<.001

[a]Δd between group effect size.

Figure 2. Means on CES-D per measurement for Chat (n=131) and WL (n=132) (EM-imputed data).



Depressive Symptoms at 7.5 Months in the Chat Group

At 7.5 months (t3), the mean CES-D score of the Chat group was 20.31 (SD 10.06) showing a mean change of 19.18 points since baseline (t_{130}=17.40, *P*<.001) and a large within group effect size of *d*=1.60 from baseline to 7.5 months. Figure 2 shows a graphical representation of CES-D outcomes.

Reliable and Clinical Change

At 9 weeks, 22.1% (29/131) participants in the Chat condition and 13.6% (18/132) in the WL condition showed reliable and clinically significant change. This difference between conditions was not significant (χ^2_1=3.24, *P*<.07). The number needed to treat was 11.7. At 4.5 months, 28.2% (37/131) in the Chat group and 11.4% (15/132) in the WL group showed a reliable and clinically significant change. This between-group difference was significant (χ^2_1=11.81, *P*<.001) and yielded a number needed to treat of 6.0. At 4.5 months, still 92 (70.2%) participants in the Chat group and 116 (87.9%) in the WL group scored 22 or higher on the CES-D, indicating they might still have clinical depression.

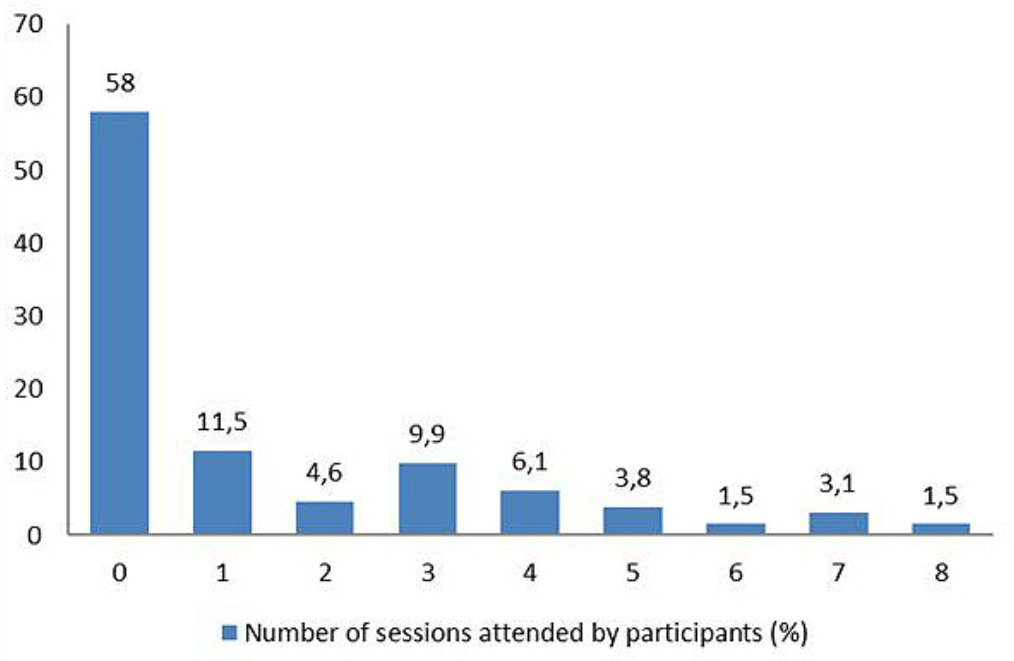
Sessions Attended and Outcome

The number of sessions attended by the subjects in the Chat condition is shown in Figure 3. The mean number of chats was 1.36 (SD 2.08), with on average 4.27 weeks (SD 6.27) between the first and last chat session (range 0-27 weeks).

According to client Web statistics, 55.7% (73/131) logged into the appointment system and 42.0% (55/131) actually had one or more chats, and 58.0% (76/131) did not have any chats. At t1 and t2, not all participants in the Chat condition had completed their therapy (n=14 had chat sessions after t1 and

n=6 after t2). There were no significant differences in changes in depressive symptoms between those who attended at least 1 chat session and those who had none (at 9 weeks: d=1.22 vs d=1.19, t_{129}=-0.16, P=.88; at 4.5 months: d=1.59 vs d=1.79, t_{129}=0.87, P=.39). When those who did not chat in the Chat condition were compared to the WL group, the non-chatters (Chat) did not differ significantly from the WL group at 9 weeks (d=1.19 vs d=1.01, $t_{187.8}$=-1.07, P=.29), but they did show better outcomes at 4.5 months (d=1.79 vs d=0.93, t_{206}=-4.50, P<.001).

Figure 3. Number of chat sessions attended by percentages of participants (n=131).



Discussion

Principal Results

The present study shows considerable improvements in depressive symptoms in both the Chat group and the waiting list group over time, but more so for the SFBT chat group, indicating it was more effective than the waiting list control condition. Between group effect sizes were small at 9 weeks (d=0.18) but increased after 4.5 months (d=0.79). At 7.5 months, the Chat condition showed further improvements. The more favorable outcomes for the Chat condition were also reflected in the significantly larger proportion of participants showing a reliable and clinically significant improvement for the Chat condition at 4.5 months, but not yet at 9 weeks. Despite the improvements, a large group had not fully recovered at 4.5 months and more than 70% of the chat intervention group still experienced depressive complaints above the cut-off (CES-D≥22) indicating they might still be struggling with depression.

Comparison With Other Work

There are no Web-based studies on SFBT interventions to compare our results with, but effect sizes reported for offline SFBT in the meta-analysis of Kim (2008) [19] are in the range of d=0.13 to d=0.26. These studies were based on only a few studies with limited numbers of participants. The study by Van der Zanden et al (2012) [12] is the most similar in target group and delivery of intervention: a chat intervention. Van der Zanden et al studied the effectiveness of a Web-based structured 6-session group chat intervention, guided by professionals, for young people aged 16-25 years, with mild to moderate depressive complaints. The intervention consisted of 6 structured sessions of CBT. As in our study, the group chat intervention proved more effective than the waiting list condition, with a large between group effect size at 3 months (d=0.94). The proportion of participants with a reliable and clinically significant change (based on the same criteria to define CES-D changes) was 56%, which is twice as high as in our study (28%). This difference in proportion of “recovered” participants might relate to the lower baseline level of depressive complaints in Van der Zanden’s study (mean 32.5 vs mean 39.6 in our study). If there are fewer depressive complaints to start with, less improvement is needed over time to reach the same threshold of 22 on the CES-D. In both studies, it was found that participants who did not chat displayed equal improvements to those who did and that non-chatters improved more at follow-up than the waiting list group. An explanation for this effect might be that not starting or discontinuing treatment could mean that participants experienced improvement and thought treatment was no longer necessary, while participants with more persistent depressive complaints started or continued treatment with hope of obtaining relief. This might especially be the case in treatments where treatment sessions are not fixed but determined by the needs of patients [41].

When compared to outcomes of face-to-face treatments for adolescents found in a meta-analytic review [42], the pre-post intervention effect size for face-to-face treatments was nearly the same (d=1.23) as the effect size found in our study at 9 weeks (d=1.20), and the between group follow-up effect size of d=0.64 of face-to-face treatments was even a bit smaller than

the effect size found in our study (d=0.79) at 4.5 months. This shows that the PratenOnline chat intervention has effects that match and even exceed those found in a meta-analysis on face-to-face treatments for adolescents.

Limitations

In our study, attrition (ie, dropping out of the study) was high. This is a phenomenon often observed in studies of Web-based interventions both among adults and youngsters [43,44]. This may have to do with the low threshold that makes it easy to start, but also easy to stop. A consequence of the high attrition rates is that a substantial number of missing observations had to be imputed, and this may have influenced outcomes. Since we do not know why attrition took place, it is hard to say if we have overestimated or underestimated the effect of the chat intervention. Results of our study, therefore, need to be considered with caution. However, the sensitivity analyses and completers only analyses show similar results to the EM-imputed data analyses, providing some confidence in the validity of the conclusions.

In our study, only 42% of those who had access to the Chat intervention of PratenOnline made use of it, although 56% had made an appointment for a chat. Limited adherence is not uncommon in Web-based interventions [45,46], and different factors can be of influence. These can be personal factors like a lack of motivation or time, an improvement or deterioration in mood, a need for face-to-face contact, or technological factors like computer problems or a lack of Internet skills [47].

In the daily chat practice of PratenOnline, the age group of 12-17 years old is highly represented. In the trial, however, only 10 (3.8%) participants in that age group were included. The major bottleneck to participate for this age group was the parental consent that had to be provided by both parents in a written consent form. Other studies also had problems with recruitment because of parental consent [12,48], making it difficult to get a handle on the effectiveness of Web-based interventions for those younger than 18 years. This also means we have to be careful in generalizing the results to young people aged 12-17 years.

In our study, a waiting list control condition was used. This might have effected the magnitude of the between group effect sizes. As Clarke et al pointed out, “the between group effect size is not just a function of the potency of the experimental intervention but is also a function of the magnitude of change observed in the control condition” (2009, page 231) [13]. Effect sizes tend to be lower when a comparison with a “strong active” intervention control condition is used [49]. But for progression of research, both studies with no-active control conditions and studies with active control conditions are necessary [13]. And although the between group effect sizes might be affected by the control group being a waiting list, the within group effects of the Chat intervention are not expected to do so that much, and these underpin the potential effects of the intervention.

Future Research Directions

As far as we know, this is the first study on a brief Web-based Solution-Focused Intervention for young people with depressive complaints. Despite the limitations of the present study, our findings indicate that adolescents and young adults with depressive symptoms can profit from access to the Web-based SFBT chat treatment. Studies in this field are few, but this one contributes to the evidence that Web-based interventions can be effective. However, because of the limitations of the study, more research is needed to find out if outcomes will be replicated. Especially for young people under the age of 18, more evidence is needed for the effectiveness of Web-based SFBT. To make such studies successful, the major impediment to include this age group needs to be tackled: the parental consent.

Acknowledgments

The study was funded by a research grant from ZonMw (Netherlands Organization for Health Research and Development), Grant No. 15700.2003. No funding bodies had any role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Also, we would like to thank all participants, without whom our study would not have been possible.

Conflicts of Interest

Pien Oijevaar was one of the founders of the chat therapy, PratenOnline, in the Netherlands, but she did not derive financial income from the PratenOnline intervention. All other authors declare no conflicts of interest.

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.2 [50].

[PDF File (Adobe PDF File), 1005KB - jmir_v16i5e141_app1.pdf]

References

1. Costello EJ, Mustillo S, Erkanli A, Keeler G, Angold A. Prevalence and development of psychiatric disorders in childhood and adolescence. Arch Gen Psychiatry 2003 Aug;60(8):837-844. [doi: 10.1001/archpsyc.60.8.837] [Medline: 12912767]
2. de Graaf R, ten Have M, van Gool C, van Dorsselaer S. Prevalence of mental disorders and trends from 1996 to 2009. Results from the Netherlands Mental Health Survey and Incidence Study-2. Soc Psychiatry Psychiatr Epidemiol 2012 Feb;47(2):203-213. [doi: 10.1007/s00127-010-0334-8] [Medline: 21197531]

3. Fergusson DM, Woodward LJ, Horwood LJ. Risk factors and life processes associated with the onset of suicidal behaviour during adolescence and early adulthood. Psychol Med 2000 Jan;30(1):23-39. [Medline: 10722173]
4. Zwaanswijk M, Van der Ende J, Verhaak PF, Bensing JM, Verhulst FC. Factors associated with adolescent mental health service need and utilization. J Am Acad Child Adolesc Psychiatry 2003 Jun;42(6):692-700. [doi: 10.1097/01.CHI.0000046862.56865.B7] [Medline: 12921477]
5. Rickwood DJ, Deane FP, Wilson CJ. When and how do young people seek professional help for mental health problems? Med J Aust 2007 Oct 1;187(7 Suppl):S35-S39. [Medline: 17908023]
6. Meredith LS, Stein BD, Paddock SM, Jaycox LH, Quinn VP, Chandra A, et al. Perceived barriers to treatment for adolescent depression. Med Care 2009 Jun;47(6):677-685. [doi: 10.1097/MLR.0b013e318190d46b] [Medline: 19434001]
7. Verhulst FC, van der Ende J. Factors associated with child mental health service use in the community. J Am Acad Child Adolesc Psychiatry 1997 Jul;36(7):901-909. [Medline: 9204667]
8. Nicholas J, Oliver K, Lee K, O'Brien M. Help-seeking and the Internet: an investigation among Australian adolescents. Australian e-Journal for the Advancement of Mental Health 2004;3(1):1-8 [FREE Full text]
9. Calear AL, Christensen H. Review of internet-based prevention and treatment programs for anxiety and depression in children and adolescents. Med J Aust 2010 Jun 7;192(11 Suppl):S12-S14. [Medline: 20528700]
10. 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]
11. 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]
12. van der Zanden R, Kramer J, Gerrits R, Cuijpers P. Effectiveness of an online group course for depression in adolescents and young adults: a randomized trial. J Med Internet Res 2012;14(3):e86 [FREE Full text] [doi: 10.2196/jmir.2033] [Medline: 22677437]
13. Clarke G, Kelleher C, Hornbrook M, Debar L, Dickerson J, Gullion C. Randomized effectiveness trial of an Internet, pure self-help, cognitive behavioral intervention for depressive symptoms in young adults. Cogn Behav Ther 2009;38(4):222-234 [FREE Full text] [doi: 10.1080/16506070802675353] [Medline: 19440896]
14. Hoek W, Schuurmans J, Koot HM, Cuijpers P. Effects of Internet-based guided self-help problem-solving therapy for adolescents with depression and anxiety: a randomized controlled trial. PLoS One 2012;7(8):e43485 [FREE Full text] [doi: 10.1371/journal.pone.0043485] [Medline: 22952691]
15. 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]
16. De Shazer S, Berg IK, Lipchik E, Nunnally E, Molnar A, Gingerich W, et al. Brief therapy: focused solution development. Fam Process 1986 Jun;25(2):207-221. [Medline: 3732502]
17. Gingerich WJ, Eisengart S. Solution-focused brief therapy: a review of the outcome research. Fam Process 2000;39(4):477-498. [Medline: 11143600]
18. Stams GJ, Dekovic M, Buist K, De Vries L. Effectiviteit van oplossingsgerichte korte therapie: een meta-analyse. Gedragstherapie 2006;39:81-94.
19. Kim JS. Examining the Effectiveness of Solution-Focused Brief Therapy: A Meta-Analysis. Research on Social Work Practice 2007 Oct 16;18(2):107-116. [doi: 10.1177/1049731507307807]
20. Corcoran J, Pillai V. A Review of the Research on Solution-Focused Therapy. British Journal of Social Work 2007 Aug 29;39(2):234-242. [doi: 10.1093/bjsw/bcm098]
21. Gingerich WJ, Peterson LT. Effectiveness of Solution-Focused Brief Therapy: A Systematic Qualitative Review of Controlled Outcome Studies. Research on Social Work Practice 2013 Jan 27:266-283. [doi: 10.1177/1049731512470859]
22. Sundstrom SM. Single-session psychotherapy for depression: Is it better to be problem-focused or solution-focused? 1993. In: Iowa State University - Thesis/Dissertation. Ames, Iowa: Iowa State University; 1993.
23. Knekt P, Lindfors O, Laaksonen MA, Renlund C, Haaramo P, Härkänen T, Helsinki Psychotherapy Study Group. Quasi-experimental study on the effectiveness of psychoanalysis, long-term and short-term psychotherapy on psychiatric symptoms, work ability and functional capacity during a 5-year follow-up. J Affect Disord 2011 Jul;132(1-2):37-47. [doi: 10.1016/j.jad.2011.01.014] [Medline: 21316768]
24. Bozeman BN. The efficacy of solution-focused therapy techniques on perceptions of hope in clients with depressive symptoms. In: New Orleans Baptist Theological Seminary - Thesis/Dissertation. Louisiana: New Orleans Baptist Theological Seminary; 1999.
25. Rhee WK, Merbaum M, Strube MJ, Self SM. Efficacy of brief telephone psychotherapy with callers to a suicide hotline. Suicide and Life-Threatening Behavior 2005;35:317-328. [doi: 10.1521/suli.2005.35.3.317]
26. Smock SA, Trepper TS, Wetchler JL, McCollum EE, Ray R, Pierce K. Solution-focused group therapy for level 1 substance abusers. J Marital Fam Ther 2008 Jan;34(1):107-120. [doi: 10.1111/j.1752-0606.2008.00056.x] [Medline: 18199184]
27. Oijevaar P. PratenOnline: een internetinterventie voor depressieve jongeren. Kinder- & Jeugdpsychotherapie 2008;35(3):1-7.

28. Cuijpers P, Boluijt P, van Straten A. Screening of depression in adolescents through the Internet : sensitivity and specificity of two screening questionnaires. Eur Child Adolesc Psychiatry 2008 Feb;17(1):32-38. [doi: 10.1007/s00787-007-0631-2] [Medline: 17876508]
29. Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, et al. The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. Biol Psychiatry 2003 Sep 1;54(5):573-583. [Medline: 12946886]
30. PratenOnline. URL: https://www.pratenonline.nl/jongeren [accessed 2014-01-20] [WebCite Cache ID 6MlPe9Ofh]
31. De Shazer S, Berg IK. 'What works?' Remarks on Research Aspects of Solution-Focused Brief Therapy. J Family Therapy 1997 May;19(2):121-124. [doi: 10.1111/1467-6427.00043]
32. EBTA Solution Focused Practice Definitions. URL: http://blog.ebta.nu/wp-content/uploads/2012/05/EBTA-SF-PRACTICE-DEFINITIONS_2012.pdf [accessed 2014-05-27] [WebCite Cache ID 6PsvFXwKe]
33. Bouma J, Ranchor AV, Sanderman R, Van Sonderen E. Het meten van symptomen van depressie met de CES-D.Een handleiding. (Measuring symptoms of depression with the CES-D. A manual). Groningen: Rijksuniversiteit Groningen; 1995.
34. Radloff LS. The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. Applied Psychological Measurement 1977 Jun 01;1(3):385-401. [doi: 10.1177/014662167700100306]
35. Ogles BM, France CR, Lunnen KM, Bell MT, Goldfarb M. Computerized depression screening and awareness. Community Ment Health J 1998 Feb;34(1):27-38. [Medline: 9559238]
36. Dempster AP, Laird NM, Rubin DB. Maximum Likelihood from Incomplete Data via the EM Algorithm. Journal of the Royal Statistical Society, Series B (Methodological) 1977;39(1):1-38 [FREE Full text]
37. Twisk JWR. Inleiding in de toegepaste biostatistiek. Maarssen: Elsevier Gezondheidszorg; 2007.
38. Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale, NJ: L. Erlbaum Associates; 1988.
39. Jacobson NS, Truax P. Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. In: Kazdin AE, editor. Methodological issues & strategies in clinical research. Washington, DC: American Psychological Association; 1992:631-648.
40. Kraemer HC, Kupfer DJ. Size of treatment effects and their importance to clinical research and practice. Biol Psychiatry 2006 Jun 1;59(11):990-996. [doi: 10.1016/j.biopsych.2005.09.014] [Medline: 16368078]
41. Barkham M, Connell J, Stiles WB, Miles JN, Margison F, Evans C, et al. Dose-effect relations and responsive regulation of treatment duration: the good enough level. J Consult Clin Psychol 2006 Feb;74(1):160-167. [doi: 10.1037/0022-006X.74.1.160] [Medline: 16551153]
42. Michael KD, Crowley SL. How effective are treatments for child and adolescent depression? A meta-analytic review. Clin Psychol Rev 2002 Mar;22(2):247-269. [Medline: 11806021]
43. 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]
44. Warmerdam L, van Straten A, Twisk J, Riper H, Cuijpers P. Internet-based treatment for adults with depressive symptoms: randomized controlled trial. J Med Internet Res 2008;10(4):e44 [FREE Full text] [doi: 10.2196/jmir.1094] [Medline: 19033149]
45. 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]
46. Melville KM, Casey LM, Kavanagh DJ. Dropout from Internet-based treatment for psychological disorders. Br J Clin Psychol 2010 Nov;49(Pt 4):455-471. [doi: 10.1348/014466509X472138] [Medline: 19799804]
47. Gerhards SAM, Abma TA, Arntz A, de Graaf LE, Evers SMAA, Huibers MJH, et al. Improving adherence and effectiveness of computerised cognitive behavioural therapy without support for depression: a qualitative study on patient experiences. J Affect Disord 2011 Mar;129(1):117-125. [doi: 10.1016/j.jad.2010.09.012] [Medline: 20889214]
48. Hoek W, Aarts F, Schuurmans J, Cuijpers P. Who are we missing? Non-participation in an Internet intervention trial for depression and anxiety in adolescents. Eur Child Adolesc Psychiatry 2012 Oct;21(10):593-595 [FREE Full text] [doi: 10.1007/s00787-012-0295-4] [Medline: 22736194]
49. Gloaguen V, Cottraux J, Cucherat M, Blackburn IM. A meta-analysis of the effects of cognitive therapy in depressed patients. J Affect Disord 1998 Apr;49(1):59-72. [Medline: 9574861]
50. 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

CES-D: Center for Epidemiologic Studies Depression Scale
CO: completers only
EM: expectation-maximization
PMM: predictive mean matching

QIDS-SR: Quick Inventory of Depressive Symptomatology—Self rated
SFBT: Solution-Focused Brief Therapy
WL: waiting list

Edited by G Eysenbach; submitted 20.01.14; peer-reviewed by A Calear, M Välimäki; comments to author 15.03.14; revised version received 09.04.14; accepted 10.05.14; published 29.05.14.

Please cite as:
Kramer J, Conijn B, Oijevaar P, Riper H
Effectiveness of a Web-Based Solution-Focused Brief Chat Treatment for Depressed Adolescents and Young Adults: Randomized Controlled Trial
J Med Internet Res 2014;16(5):e141
URL: http://www.jmir.org/2014/5/e141/
doi:10.2196/jmir.3261
PMID:24874006

Original Paper

Effectiveness of a Web-Based Tailored Intervention (E-health4Uth) and Consultation to Promote Adolescents' Health: Randomized Controlled Trial

Rienke Bannink[1], MSc; Suzanne Broeren[1], PhD; Evelien Joosten-van Zwanenburg[2], MSc; Els van As[3]; Petra van de Looij-Jansen[4], PhD; Hein Raat[1], MD, PhD

[1]Erasmus University Medical Center Rotterdam, Department of Public Health, Rotterdam, Netherlands

[2]Regional Public Health & Youth Service South-Holland South, Dordrecht, Netherlands

[3]Consortium Rivas-Careyn, Department of Youth Health Care, Gorinchem, Netherlands

[4]Municipal Public Health Service Rotterdam area, Rotterdam, Netherlands

Corresponding Author:
Hein Raat, MD, PhD
Erasmus University Medical Center Rotterdam
Department of Public Health
PO Box 2040
3000 CA
Rotterdam,
Netherlands
Phone: 31 107044634
Fax: 31 10 703 84 75
Email: h.raat@erasmusmc.nl

Abstract

Background: To promote well-being and health behaviors among adolescents, 2 interventions were implemented at 12 secondary schools. Adolescents in the E-health4Uth group received Web-based tailored messages focused on their health behaviors and well-being. Adolescents in the E-health4Uth and consultation group received the same tailored messages, but were subsequently referred to a school nurse for a consultation if they were at risk of mental health problems.

Objective: This study evaluated the effect of E-health4Uth and E-health4Uth and consultation on well-being (ie, mental health status and health-related quality of life) and health behaviors (ie, alcohol and drug use, smoking, safe sex).

Methods: A cluster randomized controlled trial was conducted among third- and fourth-year secondary school students (mean age 15.9, SD 0.69). School classes (clusters) were randomly assigned to (1) E-health4Uth group, (2) E-health4Uth and consultation group, or (3) control group (ie, care as usual). Adolescents completed a questionnaire at baseline and at 4-month follow-up assessing alcohol consumption, smoking, drug use, condom use, mental health via the Strengths and Difficulties Questionnaire (SDQ) and the Youth Self Report (YSR; only measured at follow-up), and health-related quality of life. Multilevel logistic, ordinal, and linear regression analyses were used to reveal differences in health behavior and well-being between the intervention groups and the control group at follow-up. Subsequently, it was explored whether demographics moderated the effects.

Results: Data from 1256 adolescents were analyzed. Compared to the control intervention, the E-health4Uth intervention, as a standalone intervention, showed minor positive results in health-related quality of life (B=2.79, 95% CI 0.72-4.87) and condom use during intercourse among adolescents of Dutch ethnicity (OR 3.59, 95% CI 1.71-7.55) not replicated in the E-health4Uth and consultation group. The E-health4Uth and consultation intervention showed minor positive results in the mental health status of adolescents (SDQ: B=−0.60, 95% CI −1.17 to −0.04), but a negative effect on drug use among boys (OR 0.36, 95% CI 0.13-0.96). In the subgroup of adolescents who were at risk of mental health problems at baseline (and referred for a consultation with the nurse), the E-health4Uth and consultation group showed minor to moderate positive results in mental health status (SDQ: B=−1.79, 95% CI −3.35 to −0.22; YSR: B=−9.11, 95% CI −17.52 to −0.71) and health-related quality of life (B=7.81, 95% CI 2.41-13.21) at follow-up compared to adolescents in the control group who were at risk of mental health problems at baseline.

Conclusions: Findings from this study support the use of the E-health4Uth and consultation intervention in promoting the well-being of adolescents at risk of mental health problems. Future research is needed to further evaluate the effects of the consultation as a standalone intervention, and the dual approach of further tailored eHealth messages and a consultation.

Trial Registration: Nederlands Trial Register: NTR 3596; http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=3596 (Archived by WebCite at http://www.webcitation.org/6PmgrPOuv).

(J Med Internet Res 2014;16(5):e143) doi:10.2196/jmir.3163

KEYWORDS

adolescents; youth health care; eHealth; Internet; Web-based tailoring; consultation; randomized controlled trial; health promotion; prevention

Introduction

Background

A high percentage of adolescents suffer from mental health problems, and many health-risk behaviors, such as excessive alcohol consumption, cigarette smoking, use of drugs, and having unsafe sex, are acquired during adolescence [1]. These mental health problems and health-risk behaviors often persist into adulthood, thereby affecting not only current health but also health later in life [2-4]. Given this, reducing the burden of adolescent mental health problems and health-risk behaviors is a major public health priority, one in which preventive youth health care can play an important role.

Many countries have established preventive youth health care, which refers to a variety of activities aimed at improving and protecting the health, growth, and development of young people. These activities include a system of child health care, which serves children from birth through to 18 years. In the Netherlands, all children and adolescents are invited by youth health care organizations to attend regularly scheduled preventive health consultations until the age of 13 years [5]. These consultations with a nurse or physician focus on growth, development, health functioning, and behaviors of infants, children, and adolescents. Furthermore, the consultations are funded by the government, are free of charge, and take place at the preventive youth health care center or at school. Given the rapid rate of maturation in adolescence and the mental health problems and health-risk behaviors associated with this developmental period, the government in the Netherlands encourages an additional preventive health consultation at age 15-16 years [6].

Previous research shows the use of Web-based applications for delivering tailored preventive messages in current preventive youth health care practice to be a promising development [7-9]. Web-based tailoring is a health education technique that enables the adaptation of information to individual characteristics. Web-based tailored messages eliminate (as far as possible) information that is not personally relevant [10,11] and are, therefore, more likely to be effective in changing behavior compared to nontailored messages [11]. Additionally, they facilitate the enhanced efficiency of face-to-face consultations by collecting information on adolescents' health before the consultation, which a professional can use during the consultation [7,12,13].

To promote well-being and health behaviors among adolescents, 2 interventions using Web-based tailored messages (E-health4Uth and E-health4Uth and consultation) were implemented in a preventive youth health care setting. The Web-based tailored messages focused on topics related to health-risk behaviors (eg, alcohol consumption, smoking) and well-being (eg, mental health status, suicidal thoughts). Both interventions used the same Web-based tailored messages, which were developed for adolescents (aged 12-18 years) in an earlier study [14]. In the E-health4Uth and consultation group, adolescents who were at risk of mental health problems were also referred to a school nurse for a consultation. Consequently, the intervention in this subgroup was more extensive. To facilitate communication during the consultations [7], the nurses received information regarding the adolescents' well-being and health behaviors from the E-health4Uth tool, with the adolescents' knowledge. A first investigation showed that the Web-based tailored messages and additional consultation were positively experienced by the adolescents and nurses alike [15]. However, the effectiveness of these interventions are currently unknown.

Objective of the Study

This study evaluates the effect of E-health4Uth and E-health4Uth and consultation on well-being (ie, mental health status and health-related quality of life) and health behaviors (ie, alcohol and drug use, smoking, safe sex) as applied by preventive youth health care in secondary schools. The hypotheses of the study are twofold. First, it is expected that adolescents in the E-health4Uth group will show a higher level of well-being and less risky behavior at 4-month follow-up compared to the control group (ie, care as usual). Second, it is expected that adolescents in the E-health4Uth and consultation group will show a higher level of well-being and less risky behavior (alcohol and drug use, smoking, safe sex) at 4-month follow-up compared to the control group (ie, care as usual). In addition, to gain more insight into the combined effect of E-health4Uth with a consultation, we assessed effects on well-being in the subgroup of adolescents' at risk of mental health problems at baseline, because only these adolescents were invited for a consultation with the nurse.

Methods

Study Design

A 3-armed cluster randomized controlled trial (RCT) was conducted from September 2012 to May 2013 with measurements at baseline and 4 months after the baseline measurement (trial registration: Current Controlled Trials NTR3596). The interventions were applied by preventive youth health care in secondary schools. School classes (clusters) were randomly assigned to one of the study arms (ie, E-health4Uth, E-health4Uth and consultation, control group). School classes were the unit of randomization because randomization at the

individual level (ie, the level of the adolescents) can lead to contamination of the control group [16]. A computer-generated list of random numbers was used to allocate the school classes (clusters) to one of the study arms. The randomization sequence was stratified with a 1:1:1 allocation using random block sizes of 3. This list was prepared by an investigator with no involvement in the trial and was applied by the researchers. The research proposal was reviewed by the Daily Board of the Medical Ethical Committee of Erasmus MC. As a result of this review, the Committee declared that the Medical Research Involving Human Subjects Act (also known by its Dutch abbreviation WMO) did not apply to this research proposal. The Medical Ethical Committee had no objection to the execution of this research proposal (MEC-2012-337). Further details about the study design and the interventions are provided in a design paper published elsewhere [17].

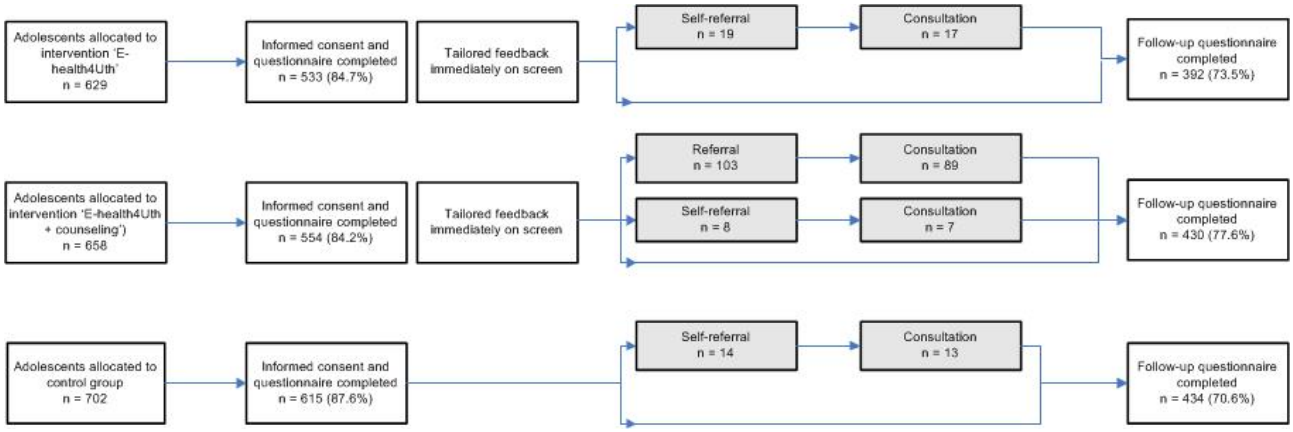
Participants and Procedures

Two youth health care organizations in the Dutch cities of Dordrecht and Zwijndrecht participated in this study and conducted the interventions in secondary schools. Of the 14 secondary schools invited by the youth health care organizations to participate in the study, 12 agreed and provided a total of 11 classes of third-year students (2 schools) and 75 classes of fourth-year students (10 schools). In the Netherlands, adolescents in the third and fourth years of secondary school are on average 15-16 years of age.

A few weeks before the start of the study, all adolescents and parents received information about the study. If parents did not want their child to participate, they could object to their child's participation. Adolescents were asked to provide written consent before they completed the baseline questionnaire. Of the 1989 eligible adolescents, 1702 (85.57%) participated: 533 (84.7%) in the E-health4Uth group, 554 (84.2%) in the E-health4Uth and consultation group, and 615 (87.6%) in the control group (Figure 1). The main reason for nonparticipation was absence, primarily because of illness. Furthermore, 29 parents refused their child's participation and 24 adolescents refused participation themselves.

At 4-month follow-up, 3 schools did not schedule the follow-up classroom assessments for all or several classes (missing data from 14 classes). At the remaining schools, 135 adolescents were absent at follow-up. In total, 1256 adolescents participated at 4-month follow-up (73.79%): 392 of 533 in the E-health4Uth group (73.5%), 430 of 554 in the E-health4Uth and consultation group (77.6%), and 434 of 615 in the control group (70.6%).

Figure 1. Flowchart of adolescents' participation.



The E-health4Uth Intervention

During one classroom session (approximately 45 min), adolescents completed a self-report questionnaire via the Internet to assess health-risk behavior and well-being with respect to the following topics: alcohol consumption, drug use, smoking, sexual behavior, bullying, mental health status, suicidal thoughts, suicide attempts, and unpleasant sexual experiences (Multimedia Appendix 1). This questionnaire served as a basis to tailor the messages, but also as a baseline measurement for the effect evaluation. The questionnaire was constructed based on several existing instruments used by municipal public health services and health institutes [18]. Consensus on the use of these instruments was established by the National Institute for Public Health and Environment (RIVM), the Dutch association for residential and homecare organizations and infant and child health clinics (Actiz), and the Association of Municipal Public Health Services in the Netherlands (GGD Nederland).

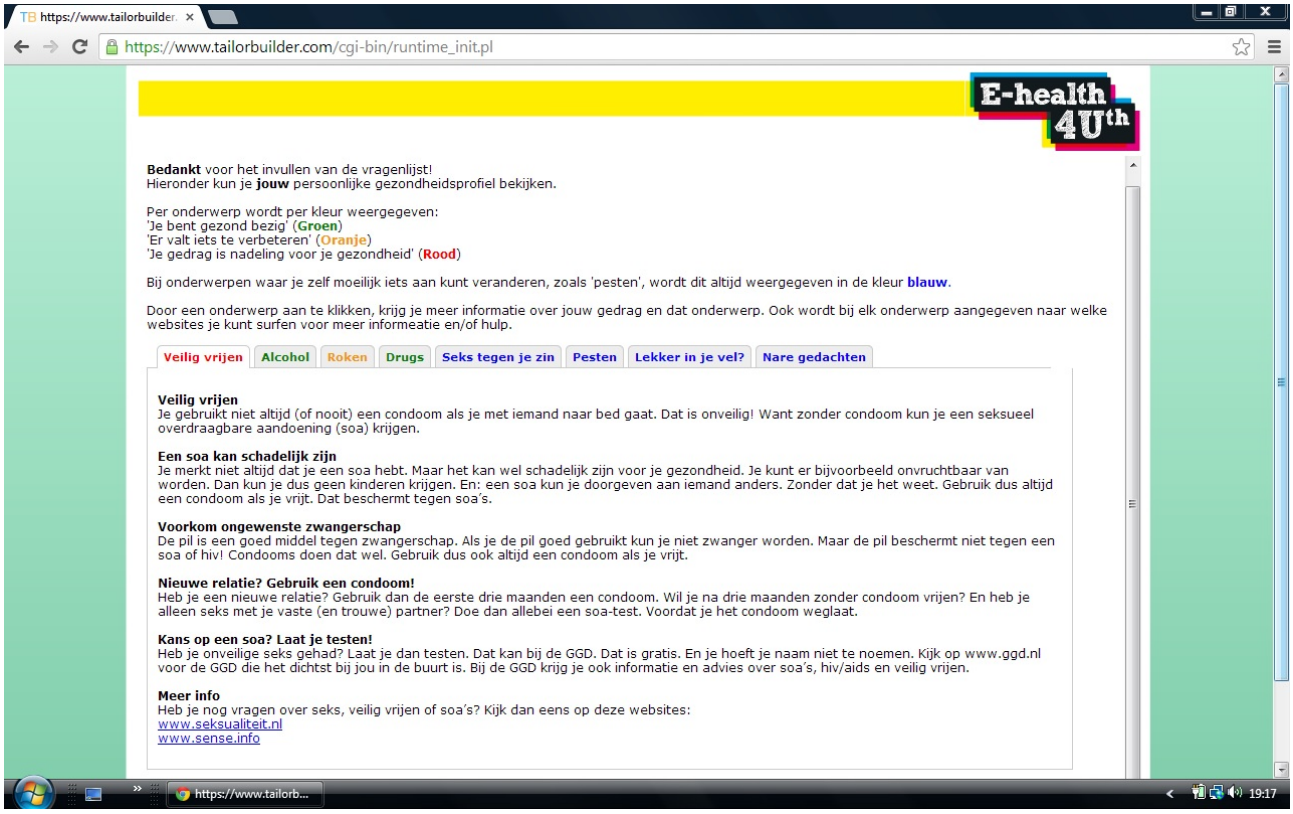
After completing the questionnaire, the participants were presented with a message of approximately the same length for each topic (see Figure 2 for an example of a message on one topic). We used Web-based tailored messages that were developed for adolescents (aged 12-18 years) and applied in an earlier study [14]. The messages were developed by the Department of Health Promotion and Health Education of the University of Maastricht. The messages were tailored to the answers given in the questionnaire. Tailored feedback is more useful in motivating people to perform the desired behaviors than nontailored feedback [11]. It also provides the opportunity to give normative feedback (ie a comparison between individual responses and the health norms) and positive feedback to reinforce desired states, both of which were used in this study.

For each topic, a score was computed which was compared with the Dutch health norms for adolescents [14,18,19]. Based on this score, a message was immediately presented on the screen that reflected the person's current behavior or well-being in relation to the Dutch health norm, and the adolescent was offered

advice to change unhealthy behavior and/or to talk to a person of trust (Figure 2). The messages were displayed in red, orange, or green, indicating unhealthy behavior, behavior just below the norm, or behavior meeting the Dutch health norm, respectively. The topics on well-being were always displayed in blue.

With links to relevant websites, adolescents were encouraged to read more information on the topics. At the end of the program, adolescents were invited to follow the Facebook page of E-health4Uth to find more information on the topics. Additionally, adolescents could check a box for a self-referral to the nurse or could send an email to the nurse. After 1 month, adolescents received a reminder of the tailored messages by email.

Figure 2. Screenshot of the Web-based tailored messages. This message was presented to adolescents who answered that they have had unsafe sex (left-most tab). The message is displayed in red, indicating unhealthy behavior. The messages on the other topics are presented when clicking on the other (colored) tabs.



The E-health4Uth and Consultation Intervention

During a classroom session, adolescents in the E-health4Uth and consultation group completed the same questionnaire and received the same intervention as that applied in the E-health4Uth--only group. Additionally, adolescents at risk of mental health problems were invited for a consultation with the nurse. Adolescents were classified as at risk of mental health problems when their score on the total problem scale of the Strengths and Difficulties Questionnaire (SDQ) was higher than 16, and/or their score on the SDQ for emotional problems was higher than 5, and/or they reported having suicidal thoughts occasionally or more frequently (or did not want to answer this question), and/or they reported a suicide attempt within the past year (or did not want to answer this question) [17].

The consultation took place at school and was provided by school nurses who were already working at the schools and who had already provided consultations to adolescents at approximately 13 years of age. These nurses were trained to apply motivational interviewing with adolescents at age 15-16 years. They received the results of the assessment for each referred adolescent before the consultation. During the consultation, the nurses focused on specific risk areas and on mental health in particular. Furthermore, they either initiated a further consultation with themselves or referred adolescents to another professional if they deemed this necessary.

Control Group

Adolescents in the control group completed the same questionnaire assessing health-risk behaviors and well-being as adolescents in the intervention groups, with the exception of the questions on unpleasant sexual experience, suicidal thoughts, and suicidal attempts, because these questions were used only to tailor the messages in the intervention group, not as measurements to assess the effectiveness of the interventions [17]. Adolescents in the control group received no messages afterwards based on their scores.

Measures

Overview

The primary outcomes of the study were adolescents' health behaviors (ie, alcohol and drug use, smoking, safe sex) and mental health status. The secondary outcome of the study was health-related quality of life. The self-report questionnaire, used

to tailor the eHealth messages in the 2 intervention groups, also served as the baseline questionnaire.

Health Behaviors

The questionnaire used to assess health behaviors was based on existing instruments previously developed by municipal public health services and health institutes in the Netherlands [18]. This questionnaire was administered at baseline and at 4-month follow-up. In this questionnaire, the frequency of alcohol consumption, smoking, drug use, and condom use were assessed on ordinal scales. Alcohol consumption was covered by 2 items: (1) how often did you drink 5 or more drinks on 1 occasion in the past 4 weeks (never to 9 or more times), and (2) how often have you been drunk or tipsy in the past 4 weeks (never to 20 or more times).

Drugs use was assessed by how often the adolescent had used drugs in the past 4 weeks (never to 20 or more times), smoking by how often the adolescent currently smokes (not at all to every day), and condom use by how often the adolescent had used condoms during intercourse (never to always). This last question was only presented if it was applicable (ie, when an adolescent had answered he/she was sexually active).

Well-Being

Mental health status was assessed by the Strengths and Difficulties Questionnaire (SDQ) [20,21] and the Youth Self Report (YSR) [22]. The SDQ consists of 25 items describing positive and negative attributes of adolescents that can be allocated to 5 subscales of 5 items each: the emotional problems, the conduct problems, the hyperactivity-inattention, the peer problems, and the prosocial behavior subscales. Each item is scored on a 3-point scale (0=not true, 1=somewhat true, and 2=certainly true). A total difficulties score is calculated by summing the scores for the emotional problems, conduct problems, hyperactivity-inattention, and peer problems subscales (range 0-40). The YSR comprises 119 items addressing emotional and behavioral problems of adolescents. Respondents have to indicate on 3-point scales to which extent each item applies to him/her (0=not, 1=sometimes, or 2=often). A total score is calculated by summing the scores on all items (range 0-210).

Health-related quality of life was measured at baseline and at 4-month follow-up by 4 items of the general health perceptions scale of the Child Health Questionnaire-Child Form (CHQ-CF-GH4) [23]. One item is scored on a 5-point scale of 1=excellent, 2=very good, 3=good, 4=moderate, 5=bad and 3 items are scored on a 5-point scale of 1=true, 2=usually true, 3=do not know, 4=usually not true, 5=not true. A total score is calculated by weighing the scores and summing the weighed scores for all items (range 0-100).

The SDQ and CHQ-CF-GH4 were administered at baseline and at 4-month follow-up. The YSR was administered only at 4-month follow-up to reduce respondent burden at baseline. At baseline, adolescents also received the E-health4Uth messages after the questionnaires.

Demographics

Age (assessed by date of birth), gender, country of birth of the adolescent and both parents, and the level of education that the adolescents attended were assessed in the baseline questionnaire. In the Netherlands, a distinction is made in the levels of education adolescents attend at secondary schools. Lower levels of education are referred to as vocational training and higher levels of education are referred to as preuniversity education. Ethnicity was classified as Dutch or non-Dutch, in accordance with the definitions of Statistics Netherlands [24]; adolescents with at least 1 parent born outside the Netherlands were classified as non-Dutch.

Statistical Analysis

Descriptive statistics were used to describe the characteristics of adolescents in the 3 study conditions. Differences between each of the intervention conditions and the control condition, as measured at baseline, were tested with independent sample *t* tests (continuous variables), Mann-Whitney *U* tests (ordinal variables), and chi-square tests (categorical variables). The effectiveness of E-health4Uth and E-health4Uth and consultation was investigated by means of multilevel logistic (categorical variables), ordinal (ordinal variable), and linear (continuous variables) regression analyses. Multilevel analysis adjusts for clusters (ie, classes) by taking the dependency between observations of adolescents from the same class into account. For the multilevel linear regression analyses, a bootstrapping method was used [25]. This method deals with data that are skewed, as is often the case with data on well-being and in this study. All regression analyses were adjusted for demographic factors that significantly differed between each of the intervention conditions and the control condition. All regression analyses were also adjusted for the baseline value of each outcome, with the exception of the YSR because this questionnaire was only assessed at follow-up. Therefore, the results of the YSR analyses were only adjusted for demographic factors that significantly differed between each of the intervention conditions and the control condition. However, exploratory analyses showed that when adjusting for the baseline value of the SDQ, which assesses a similar concept (mental health) as the YSR [21], similar results were obtained in the YSR analyses as when not adjusting for the SDQ baseline value.

To gain more insight into the combined effect of E-health4Uth with a consultation, we tested the effects on well-being in the subgroup of adolescents at risk of mental health problems at baseline because only these adolescents were referred for a consultation with the nurse. The effectiveness was tested by means of multilevel linear regression analyses (with a bootstrap procedure). The subgroup consisted of those with a score on the SDQ of >16 or a score of >5 on the emotional problems subscale of the SDQ [17]. Suicidal thoughts and suicide attempts could not be used to classify adolescents at risk of mental health problems at baseline because suicidal thoughts and suicide attempts were not assessed in the control group. These questions were not administered in this group because this group did not receive an intervention in which these concerns were addressed.

Subsequently, it was explored whether gender, ethnicity, or level of education moderated the effects of E-health4Uth and

E-health4Uth and consultation on health behaviors and well-being. This was done by adding an intervention dummy × demographic factor interaction term to the regression analyses. If the interaction term was significant at $P<.05$, a stratified analysis was conducted.

Participants were analyzed in the groups to which they had been randomized, regardless of whether they received the allocated intervention or not (eg, not attending consultation after an invitation). Each analysis of the effectiveness of the intervention was performed on the follow-up data that was available on the outcome concerned. The multilevel regression analyses were performed in Stata 13.0 (StataCorp LP, College Station, TX, USA). Other analyses were performed in SPSS 21.0 (IBM Corp, Armonk, NY, USA). The significance level was set at .05 and tests were 2-sided. To indicate the clinical significance of any benefits of the interventions, we also report odds ratios (OR) for categorical and ordinal outcomes and Cohen's *d* (*d*) for continuous outcomes.

Results

Nonresponse Analysis

Chi-square tests and *t* tests were conducted to compare adolescents participating at follow-up with adolescents not participating at follow-up. Participating at follow-up (yes/no) was used as the dependent variable and gender, age, education, ethnicity, and study condition as independent variables. Group differences were found for gender, age, education, ethnicity, and study condition, with the adolescents not participating at follow-up more often being female ($\chi^2_1=4.1$, $P=.04$), older ($t_{680}=6.69$, $P<.001$), lower educated ($\chi^2_1=20.0$, $P<.001$), of non-Dutch ethnicity ($\chi^2_1=64.7$, $P<.001$), and allocated to the control group instead of the E-health4Uth and consultation group ($\chi^2_1=7.5$, $P=.006$).

Adolescents' Characteristics

The average age of the adolescents in this study was 15.9 years (SD 0.69); 54.70% (687/1256) of the sample consisted of boys, 76.19% (957/1256 were of Dutch ethnicity, 50.48% (634/1256) attended vocational training, and 49.52% (622/1256) preuniversity education. Table 1 shows general characteristics and baseline health behaviors and well-being of adolescents in the 3 study conditions. At baseline, a lower percentage of adolescents in the E-health4Uth group had used drugs in the past 4 weeks compared to adolescents in the control group (4.6% vs 8.1%; $P<.04$). Further, adolescents in the E-health4Uth and consultation group were significantly younger than adolescents in the control group (mean 15.95, SD 0.70 vs mean 15.79, SD 0.66; $P<.001$). Therefore, all analyses evaluating the effectiveness of E-health4Uth and consultation were adjusted for age.

Table 1. General characteristics and baseline health behaviors and well-being of adolescents for the intervention groups and control group (N=1256).

Characteristics	E-health4Uth n=392	E-health4Uth + consult n=430	Control n=434	*P* value	
				E-health4Uth vs control	E-health4Uth + consult vs control
Number of school classes	27	26	25		
Age (years), mean (SD)	15.84 (0.70)	15.95 (0.70)	15.79 (0.66)	.28	<.001[a]
Gender, n (%)					
Male	223 (56.9)	241 (56.0)	223 (51.4)	.11	.17[b]
Female	169 (43.1)	189 (44.0)	211 (48.6)		
Ethnicity, n (%)					
Dutch	311 (79.3)	320 (74.4)	326 (75.1)	.15	.81[b]
Non-Dutch	81 (20.7)	110 (25.6)	108 (24.9)		
Educational level, n (%)					
Vocational training	191 (48.7)	231 (53.7)	212 (48.8)	.97	.15[b]
Preuniversity	201 (51.3)	199 (46.3)	222 (51.2)		
Alcohol consumption					
5 or more drinks on 1 occasion in the past 4 weeks, n (%)					
0 times	255 (65.1)	272 (63.4)	292 (67.6)	.48[c]	.20[c]
1 times	962 (15.8)	69 (16.1)	62 (14.4)		
2 times	36 (9.2)	36 (8.4)	34 (7.9)		
3-4 times	22 (5.6)	35 (8.2)	29 (6.7)		
5 or more times	17 (4.3)	17 (4.0)	15 (3.5)		
Have been drunk or tipsy in the past 4 weeks, n (%)					
0 times	290 (74.0)	318 (74.1)	333 (77.1)	.28[c]	.30[c]
1 times	54 (13.8)	60 (14.0)	53 (12.3)		
2 times	21 (5.4)	22 (5.1)	24 (5.6)		
3 or more times	27 (6.9)	29 (6.8)	22 (5.1)		
Smoking, n (%)					
No	329 (83.9)	352 (82.1)	352 (81.5)	.39[c]	.92[c]
Less than once a week	13 (3.3)	16 (3.7)	19 (4.4)		
At least once a week, but not every day	15 (3.8)	14 (3.3)	20 (4.6)		
Every day	35 (8.9)	47 (11.0)	41 (9.5)		
Drug use (past 4 weeks), n (%)					
0 times	274 (95.4)	402 (93.7)	397 (91.9)	.04[b]	.31[b]
1 or more times	18 (4.6)	27 (6.3)	35 (8.1)		
Condom use during intercourse (n=324), n (%)					
Always	52 (53.1)	68 (52.3)	49 (51.0)	.50[c]	.55[c]
Usually	21 (21.4)	25 (19.2)	15 (15.6)		
Sometimes/almost never	14 (14.3)	25 (19.2)	18 (18.8)		
Never	11 (11.2)	12 (9.2)	14.6 (14)		
Well-being, mean (SD) [d]					

Characteristics	E-health4Uth n=392	E-health4Uth + consult n=430	Control n=434	*P* value	
				E-health4Uth vs control	E-health4Uth + consult vs control
SDQ score	10.06 (5.57)	9.75 (5.14)	9.91 (5.32)	.69[a]	.67[a]
CHQ-CF-GH4 score	71.39 (17.87)	71.62 (18.49)	73.67 (17.78)	.07[a]	.10[a]

[a]Independent-samples *t* test.

[b]Chi-square test.

[c]Mann-Whitney *U* test.

[d]SDQ: Strengths and Difficulties Questionnaire (higher score indicates more mental health problems; range 0-40); CHQ-CF-GH4: Child Health Questionnaire-Child Form-General Health (higher score indicates a better health-related quality of life; range 0-100).

Table 2. Follow-up health behaviors and well-being of adolescents and effects of the interventions with the control group as reference (N=1252).

Behaviors and well-being	E-health4Uth n=392	E-health4Uth + consult n=430	Control n=434	E-health4Uth vs control[a]		E-health4Uth + consult vs control group[b]	
				OR/B (95% CI)	*P*	OR/B (95% CI)	*P*
Alcohol consumption, n (%)							
5 or more drinks on 1 occasion in the past 4 weeks							
0 times	230 (59.0)	280 (65.9)	276 (63.7)	0.90 (0.61, 1.34)[c]	.62	1.21 (0.77, 1.26)[c]	.35
1 times	62 (15.9)	44 (10.4)	58 (13.4)				
2 times	43 (11.0)	32 (7.5)	37 (8.5)				
3-4 times	28 (7.2)	46 (10.8)	34 (7.9)				
5 or more times	27 (6.9)	23 (5.4)	28 (6.5)				
Have been drunk or tipsy in the past 4 weeks							
0 times	275 (70.5)	317 (74.6)	321 (74.1)	0.90 (0.61, 1.35)[c]	.62	1.22 (0.85, 1.74)[c]	.29
1 times	57 (14.6)	52 (12.2)	57 (13.2)				
2 times	18 (4.6)	20 (4.7)	20 (4.6)				
3 or more times	40 (10.3)	36 (8.5)	35 (8.1)				
Smoking, n (%)							
No	323 (82.8)	351 (82.6)	349 (80.8)	0.97 (0.61, 1.56)[c]	.90	0.95 (0.58, 1.57)[c]	.84
Less than once a week	21 (5.4)	14 (3.3)	23 (5.3)				
At least once a week, but not every day	14 (3.6)	11 (2.6)	19 (4.4)				
Every day	32 (8.2)	49 (11.5)	41 (9.5)				
Drug use (past 4 weeks), n (%)							
0 times	367 (94.1)	381 (89.6)	396 (91.7)	1.06 (0.43, 2.61)[d]	.90	0.65 (0.26, 1.59)[d]	.34
1 or more times	23 (5.9)	44 (10.4)	36 (8.3)				
Condom use during intercourse (n=376), n (%)							
Always	62 (52.1)	66 (43.7)	43 (40.6)	2.09 (1.04, 4.22)[c]	.04	1.36 (0.76, 2.44)[c]	.31
Usually	24 (20.2)	32 (21.2)	15 (14.2)				
Sometimes/almost never	18 (15.1)	38 (25.2)	27 (25.5)				
Never	15 (12.6) (15)	15 (9.9)	21 (19.8)				
Well-being, mean (SD)							
SDQ score[f]	8.92 (5.26)	8.42 (5.05)	9.07 (5.38)	−0.24 (−0.78, 0.29)[e]	.37	−0.60 (−1.17, −0.04)[e]	.04
YSR score[f]	33.89 (23.02)	31.58 (22.58)	34.75 (25.26)	−0.89 (−4.18, 2.40)[e]	.60	−2.74 (−5.92, 0.44)[e]	.09
CHQ-CF-GH4 score[g]	75.34 (16.56)	74.00 (18.49)	73.73 (18.17)	2.79 (0.72, 4.87)[e]	.008	1.03 (−1.12, 3.19)[e]	.35

[a]E-health4Uth vs control group: analyses were adjusted for the baseline value of each outcome.

[b]E-health4Uth and consultation vs control group: analyses were adjusted for age and the baseline value of each outcome.

[c]Multilevel ordinal regression; OR (95% CI).

[d]Multilevel logistic regression; OR (95% CI).

[e]Multilevel linear regression; Beta coefficient (95% CI).

[f]A higher score indicates more mental health problems (SDQ range 0-40, YSR range 0-210).

[g]A higher score indicates a better health-related quality of life (range 0-100).

Effects of E-health4Uth

Adolescents in the E-health4Uth group used condoms significantly more often at follow-up compared to adolescents in the control group (52.1% vs 40.6%; OR 2.09, 95% CI 1.04-4.22) (Table 2). Furthermore, the health-related quality of life of adolescents in the E-health4Uth group was significantly better at follow-up compared to adolescents in the control group (mean 75.34, SD 16.56 vs mean 73.73, SD 18.17; B=2.79, 95% CI 0.72-4.87; *d*=0.09). No other effects of the E-health4Uth intervention on health behaviors or well-being were found.

Effects of E-health4Uth and consultation

At follow-up, adolescents in the E-health4Uth and consultation group reported a significantly better mental health status compared to adolescents in the control group (SDQ: mean 8.42, SD 5.05 vs 9.07, SD 5.38; B=−0.60, 95% CI −1.17 to −0.04; *d*=0.12) (Table 2). No effects of the E-health4Uth and consultation intervention on health behaviors were found.

Adolescents in the E-health4Uth and consultation group, who were at risk of mental health problems at baseline and were therefore referred for a consultation with the nurse, reported a significantly better mental health status (SDQ: mean 12.79, SD 5.63 vs 14.57, SD 5.03; B=−1.79, 95% CI −3.35 to −0.22; *d*=0.33; YSR: mean 48.13, SD 25.45 vs 57.12, SD 27.66; B=−9.11, 95% CI −17.52 to −0.71; *d*=0.34) and a better health-related quality of life (mean 69.56, SD 18.37 vs 62.53, SD 20.08; B=7.81, 95% CI 2.41-13.21; *d*=0.37) at follow-up compared to adolescents in the control group who were at risk of mental health problems at baseline (Table 3). These results were not replicated among adolescents who were at risk of mental health problems in the E-health4Uth standalone intervention group (Table 3), indicating that the dual approach of advice and a consultation (ie, E-health4Uth and consultation) may have been responsible for the positive effects on well-being.

Table 3. Follow-up well-being of adolescents who were at risk of mental health problems at baseline and the effects of the interventions on well-being, with the control group as reference (n=194).

Well-being	Adolescents at risk of mental health problems, mean (SD)			E-health4Uth vs control[b]		E-health4Uth + consult vs control[c]	
	E-health4Uth[a] n=63	E-health4Uth + consult[a] n=63	Control[a] n=68	Beta (95% CI)	*P*	Beta (95% CI)	*P*
SDQ score[d]	14.44 (5.67)	12.79 (5.63)	14.57 (5.03)	0.04 (−1.60, 1.68)	.96	−1.79 (−3.35, −0.22)	.03
YSR score[d]	56.49 (27.86)	48.13 (25.45)	57.12 (27.66)	−0.63 (−9.72, 8.47)	.89	−9.11 (−17.52, −0.71)	.03
CHQ-CF-GH4[e]	67.59 (17.14)	69.56 (18.37)	62.53 (20.08)	4.78 (−0.70, 10.25)	.09	7.81 (2.41, 13.21)	.005

[a]In the E-health4Uth (5 of 63) and control group (4 of 68), some adolescents at risk of mental health problems also attended the consultation after they referred themselves to the nurse. In the E-health4Uth and consultation group, 57 of the 63 referred adolescents attended the consultation.

[b]E-health4Uth vs control group: analyses were adjusted for the baseline value of each outcome. Multilevel linear regression.

[c]E-health4Uth and consultation vs control group: analyses were adjusted for age and the baseline value of each outcome. Multilevel linear regression.

[d]A higher score indicates more mental health problems (SDQ range 0-40, YSR range 0-210).

[e]A higher score indicates a better health-related quality of life (range 0-100).

Interaction Effects

Exploratory interaction analyses showed 3 statistically significant interactions between the dummy variables for intervention groups and the demographic factors. Ethnicity moderated the intervention effect of E-health4Uth on condom use and gender moderated the intervention effect of E-health4Uth on alcohol consumption and the intervention effect of E-health4Uth and consultation on drug use (Table 4). More specifically, adolescents of Dutch ethnicity in the E-health4Uth group were more likely to use condoms during intercourse at follow-up compared to adolescents of Dutch ethnicity in the control group (OR 3.59, 95% CI 1.71-7.55), whereas there was no significant effect of the intervention among adolescents of non-Dutch ethnicity (OR 0.25, 95% CI 0.02-2.49). Furthermore, boys in the E-health4Uth and consultation group were more likely to use drugs at follow-up compared to boys in the control group (OR 0.36, 95% CI 0.13-0.96), whereas there was no significant intervention effect among girls (OR 4.47, 95% CI 0.72-27.74). Among boys and girls in the E-health4Uth and consultation group, no significant intervention effect was found on alcohol consumption (boys: OR 0.68, 95% CI 0.40-1.15, girls: OR 1.35, 95% CI 0.76-2.38).

Table 4. Stratified analyses of intervention effects on health behavior for the various levels of the significant moderator variables.[a]

Outcome	E-health4Uth vs control group		E-health4Uth +consult vs control group	
	OR (95% CI)[b]	*P* value	OR (95% CI)[c]	*P* value
Alcohol consumption				
Have been drunk or tipsy in the past 4 weeks				
Gender				
Boys	0.68 (0.40, 1.15)	.15		
Girls	1.35 (0.76, 2.38)	.31		
Drugs use (past 4 weeks)				
Gender				
Boys			0.36 (0.13, 0.96)	.04
Girls			4.47 (0.72, 27.74)	.11
Condom use during intercourse				
Ethnicity				
Dutch	3.59 (1.71, 7.55)	.001		
Non-Dutch	0.25 (0.03, 2.49)	.38		

[a]Only the results of the stratified analyses according to the significant moderators of the intervention effects are presented.

[b]Multilevel ordinal regression.

[c]Multilevel logistic regression.

Discussion

Principal Results

Using a cluster RCT, we evaluated the effect of E-health4Uth as a standalone intervention and in combination with an additional consultation for adolescents who were at risk of mental health problems. The E-health4Uth intervention as a standalone intervention showed minor positive results in a small number of outcomes, namely in the health-related quality of life and in condom use during intercourse among adolescents of Dutch ethnicity. The 2 positive results found in the E-health4Uth intervention were not replicated in the E-health4Uth and consultation group. The E-health4Uth and consultation intervention showed minor positive results in the mental health status of adolescents, but a negative effect on drug use among boys was found. In the subgroup of adolescents who were at risk of mental health problems at baseline and were, therefore, referred for a consultation with the nurse, the E-health4Uth and consultation group showed small to moderate positive results on mental health status and health-related quality of life at follow-up compared to adolescents in the control group who were at risk of mental health problems at baseline.

Interpretation

Although it is promising that positive effects were found in the E-health4Uth group, only a small number of outcome measures were statistically significant (ie, health-related quality of life and condom use during intercourse), the effects were small, and the effects on condom use were only found among adolescents of Dutch ethnicity. Furthermore, because the E-health4Uth and consultation group received the same messages as the E-health4Uth group plus an additional consultation for the adolescents at risk of mental health problems, one would expect that the effects on condom use and health-related quality of life would have also been present in the E-health4Uth and consultation group. Although these effects pointed in the same direction, they were not significant in the E-health4Uth and consultation group. Therefore, the effects found in the E-health4Uth group have to be interpreted with caution.

In contrast to our hypothesis, we could not demonstrate that the E-health4Uth intervention was effective in promoting other health behaviors or the mental health status of adolescents. Although various studies show that Web-based tailoring is a promising technique to promote health behaviors and mental health status of adolescents [26-34], most studies are focused on older adolescents. Furthermore, the results of the evaluation of the appreciation of the tailored messages used in this trial showed that adolescents did not evaluate the tailored messages as explicitly positive in terms of their personally relevance [15]. If messages are not deemed personally relevant, the positive effect of these messages may be reduced [35]. Therefore, the tailored messages used in this study could potentially be improved further, possible resulting in messages that are more personally relevant and effective. The current messages could be further tailored by using, for example, demographics, personal cognitive factors (eg, manner in which health risks are perceived by the individual), social factors (eg, susceptibility to peer pressure), or the self-efficacy of the individual (eg, judgment of capability to change unhealthy behavior) [36,37]. Furthermore, algorithms generating tailored information can be easily extended to use more characteristics of the adolescent to tailor the messages, whereas wide-scale distribution can be arranged at relatively low cost.

Moreover, knowledge on how adolescents process and respond to personalized feedback is currently scarce [38]. More insight into how adolescents process the feedback messages, single messages, and when receiving multiple feedback messages on various behaviors at one point in time is needed to be able to improve interventions. Although the focus on multiple behaviors is becoming an increasingly popular strategy in interventions using Web-based tailored messages [39-41], adolescents receive a lot of information at the same time and it is conceivable that adolescents become overwhelmed by the amount of information. Furthermore, tailored messages were used and appreciated positively by adolescents in this trial [15] and they seemed interested in receiving feedback on health behaviors. In contrast to older people who are confronted with chronic diseases more often, adolescents are probably less likely to see the benefits of health behavior changes and consequently less likely to be internally motivated to invest time in health behavior changes [42].

As hypothesized, the E-health4Uth and consultation intervention was effective in enhancing the mental health status of adolescents. Furthermore, it is promising that expanding the Web-based tailored intervention with a consultation in the subgroup of adolescents at risk of mental health problems, improved the effectiveness of the intervention on mental health and health-related quality of life among these adolescents. The effect of the E-health4Uth and consultation intervention on the well-being of adolescents at risk of mental health problems was minor to moderate, in-line with the results of previous studies in which adolescents at risk of depression and anxiety, 2 components of the broader construct of mental health, benefited from an Internet program combined with a consultation [43-46]. A potential explanation for the effects on the well-being of adolescents is the dual approach of advice and a consultation. This approach guaranteed a repetition of the main mental health message and combined digital and oral feedback. However, it is also feasible that the consultation was responsible for the positive effects that were found and that the E-health4Uth questionnaire was primarily a useful way to select adolescents who needed further face-to-face support.

Because the nurses rated the information they received about the adolescents before the consultation as helpful in most cases (80.0%) [15], this information on adolescents' health may have supported the nurse during the consultation to better tailor the information provided to the adolescent's needs, thereby enhancing the effectiveness.

In contrast to our hypothesis, positive effects in the E-health4Uth and consultation group were not found in promoting health behaviors. Therefore, it might also be beneficial to apply the dual approach of advice and a consultation to the health behavior messages (ie, expand the Web-based tailored messages on the health behaviors with a consultation focused on these health behaviors), instead of primarily focusing on mental health in the consultation. A previous study, integrating Web-based tailored messages on fruit and vegetable intake with a consultation focused on this intake among schoolchildren showed promising results in the area of preventive youth health care [35]. However, future research is needed to investigate the degree to which the impact of Web-based tailored messages on health behaviors may be enhanced through expanding these messages with a consultation. Especially since an unexpected negative effect on drug use among boys was found in the E-health4Uth and consultation group. Although this result could be a random effect, another possible explanation is that giving information about drug use to adolescents raises adolescents' curiosity about trying drugs. In earlier studies, a similar negative effect on drug use among Dutch adolescents was found [47,48]. In one of these studies, it was found that this increase in frequency was only a temporary effect [48]. However, it is an indication that one has to be careful with health promotion on drug use among adolescents and it highlights the importance of careful evaluation and in-depth study of how health promotion on drug use works for adolescents.

Because the Web-based tailored messages and the additional consultation were already interwoven with the existing practice of preventive youth health care, they are especially promising for future implementation. Implementing the Web-based tailored messages as a universal program (ie, offering it to all adolescents in a school class regardless of current symptom level or risk status) has multiple benefits. Universal programs, instead of programs that only focus on adolescents who are at risk (eg, for mental health problems), are often preferred by school administrators [49]. Additionally, by collecting information on the health of all adolescents in a school class, this approach presents an opportunity to select vulnerable adolescents and to enhance the efficiency of face-to-face consultations [7,12,13]. Efficiency is essential given the current financial strain on preventive health care.

Strengths and Limitations

Important strengths of this study are the randomized controlled design and large sample size. The response rate was relatively high and our study population resembles the average Dutch adolescent population in secondary schools for gender, ethnicity, and education level [50]. However, this study was conducted only among Dutch adolescents aged 15-16 years in a preventive care setting; therefore, generalization to other countries, age groups, and settings should be done with caution. Furthermore, dropout was higher among girls, older adolescents, adolescents with a low education level, adolescents of non-Dutch ethnicity, and adolescents allocated to the control group instead of the E-health4Uth and consultation group, which could also affect the generalizability of the results. Nevertheless, we expect that the effects of our study would have been stronger without this selective dropout. A vulnerable group research has shown to be at a particularly heightened risk of mental health problems [15] and of exhibiting unhealthy behavior [1,51], dropped out while interventions are especially effective in high risk groups [52].

Additionally, the use of self-report measures may have resulted in less reliable outcomes. Therefore, the collection of more objective data on health behavior and additional parent and teacher ratings on the well-being of the adolescents may have been useful. Nevertheless, research suggests, for example, that self-reported alcohol consumption among adolescents is generally considered valid [53] and that adolescents are better reporters of their own mental health status than parents and teachers [54]. The percentages of adolescents with unhealthy

behaviors and mental health problems in this study are largely comparable with the percentages of adolescents aged 15-16 years with unhealthy behaviors and mental health problems in the Netherlands [1]. Another limitation is that some adolescents in the control group may have received information from friends in the intervention groups despite the randomization of school classes. This may have contaminated the results. Furthermore, the overlap between the 2 intervention groups and control group is a limitation. In the E-health4Uth and consultation group, only adolescents at risk of mental health problems were invited for a consultation; thus, the other adolescents in this group received actually only the E-health4Uth intervention. Moreover, adolescents in all the groups could ask for a self-referral with the school nurse. Although only a few adolescents in the E-health4Uth group (17 of 533) and the control group (13 of 615) attended a consultation with the nurse (Figure 1), this may have underestimated the results. Therefore, in addition to the intention-to-treat analyses in which we analyzed adolescents in the groups to which they were randomized, regardless of whether they received the allocated intervention or not (eg, whether they attended the consultation or not after an invitation), we conducted exploratory per-protocol analyses. For these exploratory analyses, adolescents were allocated to the intervention they actually received (eg, the adolescents who self-referred to the nurse for a consultation were included in the E-health4Uth and consultation group for analysis purposes). The results from these per-protocol analyses were stronger than the results from the intention-to-treat analyses. That is, these analyses showed larger effects on mental health and health-related quality of life for the subgroup of adolescents at risk of mental health problems at baseline than the intention-to-treat analyses, suggesting that the results presented in this study may be underestimations of the actual effects. Unfortunately, information about the percentages of adolescents invited for a further consultation with the nurse or who were referred to another professional was not available because this information was not consistently administered by the nurse. However, the available data suggest that a low percentage of adolescents were invited for a further consultation with the nurse or referred to another professional. Further research is necessary to assess whether the consultation is an effective way in selecting adolescents who need help and providing them with the help they need.

Conclusions

Findings from this study support the use of the E-health4Uth and consultation intervention in promoting the well-being of adolescents at risk of mental health problems. Compared to care as usual, E-health4Uth combined with a consultation was effective in promoting the mental health status and health-related quality of life in the subgroup of adolescents at risk of mental health problems. It is feasible that the consultation (and not the dual approach) was primarily responsible for these positive effects. However, E-health4Uth may have been a valuable tool to select vulnerable adolescents and to provide the nurse with information about the health of these adolescents. This could have contributed to the efficiency of the face-to-face consultation. Because the E-health4Uth and consultation intervention can be embedded in the existing practice of preventive youth health care, this increases the chance of future implementation. However, more research is needed to further evaluate the effects of the consultation as a standalone intervention and of the dual approach of further tailored eHealth messages and a consultation. Adding a consultation for adolescents at risk of mental health problems seems promising; therefore, future research is recommended to evaluate the potential effect of a consultation for adolescents who exhibit unhealthy behavior.

Acknowledgments

The authors thank the adolescents and the staff of the preventive youth health care centers in Dordrecht and Zwijndrecht for their participation in this study. This work was supported by the Netherlands Organisation for Health Research and Development (ZonMw) (grant number 156511010). The publication of this study was supported by a grant of the Netherlands Organisation for Scientific Research (NWO).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Topics of the E-health4Uth modules.

[PDF File (Adobe PDF File), 27KB - jmir_v16i5e143_app1.pdf]

Multimedia Appendix 2

CONSORT-EHEALTH checklist V1.6.2 [55].

[PDF File (Adobe PDF File), 670KB - jmir_v16i5e143_app2.pdf]

References

1. van Dorsselaer S, de Looze M, Vermeulen-Smit E, de Roos S, Verdurmen J, ter Bogt T. Gezondheid, welzijn en opvoeding van jongeren in nederland. Utrecht: Trimbos-instituut, Universiteit Utrecht, Sociaal en cultureel planbureau; 2009.
2. Fergusson DM, Woodward LJ. Mental health, educational, and social role outcomes of adolescents with depression. Arch Gen Psychiatry 2002 Mar;59(3):225-231. [Medline: 11879160]
3. Hofstra MB, van der Ende J, Verhulst FC. Child and adolescent problems predict DSM-IV disorders in adulthood: a 14-year follow-up of a Dutch epidemiological sample. J Am Acad Child Adolesc Psychiatry 2002 Feb;41(2):182-189. [doi: 10.1097/00004583-200202000-00012] [Medline: 11837408]
4. DeWit DJ, Adlaf EM, Offord DR, Ogborne AC. Age at first alcohol use: a risk factor for the development of alcohol disorders. Am J Psychiatry 2000 May;157(5):745-750. [Medline: 10784467]
5. Ministerie van Volksgezondheid, Welzijn en Sport. Basistakenpakket jeugdgezondheidszorg 0-19 jaar. Den Haag: Ministerie van Volksgezondheid, Welzijn en Sport; 2002.
6. van Heerwaarden Y. De jgz in beeld bij adolescenten. Samen bouwen aan gezondheid en gezond gedrag voor duurzame participatie van jongeren. Utrecht: Nederlands Centrum Jeugdgezondheidszorg (NCJ); 2013.
7. Mangunkusumo R, Brug J, Duisterhout J, de Koning H, Raat H. Feasibility, acceptability, and quality of Internet-administered adolescent health promotion in a preventive-care setting. Health Educ Res 2007 Feb;22(1):1-13 [FREE Full text] [doi: 10.1093/her/cyl010] [Medline: 16714366]
8. van Beelen ME, Vogel I, Beirens TM, Kloek GC, den Hertog P, van der Veen MD, et al. Web-Based eHealth to Support Counseling in Routine Well-Child Care: Pilot Study of E-health4Uth Home Safety. JMIR Res Protoc 2013;2(1):e9 [FREE Full text] [doi: 10.2196/resprot.1862] [Medline: 23611794]
9. van Beelen MEJ, Beirens TMJ, van Beeck EF, Raat H. Effectiveness of web-based, tailored advice on parents' child safety behaviours: A randomized controlled trial. J Med Internet Res 2014 Jan 24:e-17. [doi: 10.2196/jmir.2521]
10. Kreuter MW, Farell D, Olevitch L, Brennan L. Tailoring Health Messages: Customizing Communication with Computer Technology. London: Lawrence Erlbaum Associates; 2000.
11. Kroeze W, Werkman A, Brug J. A systematic review of randomized trials on the effectiveness of computer-tailored education on physical activity and dietary behaviors. Ann Behav Med 2006 Jun;31(3):205-223. [doi: 10.1207/s15324796abm3103_2] [Medline: 16700634]
12. Paperny DM, Hedberg VA. Computer-assisted health counselor visits: a low-cost model for comprehensive adolescent preventive services. Arch Pediatr Adolesc Med 1999 Jan;153(1):63-67. [Medline: 9895001]
13. Sciamanna CN, Novak SP, Houston TK, Gramling R, Marcus BH. Visit satisfaction and tailored health behavior communications in primary care. Am J Prev Med 2004 Jun;26(5):426-430. [doi: 10.1016/j.amepre.2004.02.003] [Medline: 15165659]
14. de Nooijer J, de Vries NK. Monitoring health risk behavior of Dutch adolescents and the development of health promoting policies and activities: the E-MOVO project. Health Promot Int 2007 Mar;22(1):5-10 [FREE Full text] [doi: 10.1093/heapro/dal036] [Medline: 17015406]
15. Bannink R, Broeren S, Joosten-van Zwanenburg E, Van As E, Van de Looij-Jansen P, Raat H. Use and appreciation of a web-based, tailored intervention (E-health4uth) combined with counseling to promote adolescents' health in preventive youth health care. JMIR Res Protoc. In press . [doi: 10.2196/resprot.2855]
16. Campbell MK, Elbourne DR, Altman DG, CONSORT group. CONSORT statement: extension to cluster randomised trials. BMJ 2004 Mar 20;328(7441):702-708 [FREE Full text] [doi: 10.1136/bmj.328.7441.702] [Medline: 15031246]
17. Bannink R, Joosten-van Zwanenburg E, van de Looij-Jansen P, van As E, Raat H. Evaluation of computer-tailored health education ('E-health4Uth') combined with personal counselling ('E-health4Uth + counselling') on adolescents' behaviours and mental health status: design of a three-armed cluster randomised controlled trial. BMC Public Health 2012;12:1083 [FREE Full text] [doi: 10.1186/1471-2458-12-1083] [Medline: 23244449]
18. Monitor gezondheid. Lokale en nationale monitor gezondheid. 2011. URL: https://www.monitorgezondheid.nl/ [accessed 2013-07-30] [WebCite Cache ID 6IUiOeXYv]
19. de Nooijer J, Veling ML, Ton A, de Vries H, de Vries NK. Electronic monitoring and health promotion: an evaluation of the E-MOVO Web site by adolescents. Health Educ Res 2008 Jun;23(3):382-391 [FREE Full text] [doi: 10.1093/her/cym086] [Medline: 18187490]
20. Goodman R, Ford T, Simmons H, Gatward R, Meltzer H. Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. Br J Psychiatry 2000 Dec;177:534-539 [FREE Full text] [Medline: 11102329]
21. Janssens A, Deboutte D. Screening for psychopathology in child welfare: the Strengths and Difficulties Questionnaire (SDQ) compared with the Achenbach System of Empirically Based Assessment (ASEBA). Eur Child Adolesc Psychiatry 2009 Nov;18(11):691-700. [doi: 10.1007/s00787-009-0030-y] [Medline: 19462154]
22. Achenbach TM, Rescorla LA. Manual for the ASEBA School-Age Forms & Profiles: An Integrated System of Multi-Informant Assessment. Burlington, VT: ASEBA; 2001.
23. Raat H, Landgraf JM, Bonsel GJ, Gemke RJ, Essink-Bot ML. Reliability and validity of the child health questionnaire-child form (CHQ-CF87) in a Dutch adolescent population. Qual Life Res 2002 Sep;11(6):575-581. [Medline: 12206578]

24. Centraal bureau voor de statistiek. Allochtoon. 2013. URL: http://www.cbs.nl/nl-NL/menu/methoden/begrippen/default.htm?ConceptID=37 [accessed 2013-10-29] [WebCite Cache ID 6KjW8ULzH]
25. Fox J. Applied Regression Analysis and Generalized Linear Models. Los Angeles: Sage; 2008.
26. Kiene SM, Barta WD. A brief individualized computer-delivered sexual risk reduction intervention increases HIV/AIDS preventive behavior. J Adolesc Health 2006 Sep;39(3):404-410. [doi: 10.1016/j.jadohealth.2005.12.029] [Medline: 16919803]
27. Spijkerman R, Roek MA, Vermulst A, Lemmers L, Huiberts A, Engels RC. Effectiveness of a web-based brief alcohol intervention and added value of normative feedback in reducing underage drinking: a randomized controlled trial. J Med Internet Res 2010;12(5):e65 [FREE Full text] [doi: 10.2196/jmir.1465] [Medline: 21169172]
28. Hustad JT, Barnett NP, Borsari B, Jackson KM. Web-based alcohol prevention for incoming college students: a randomized controlled trial. Addict Behav 2010 Mar;35(3):183-189 [FREE Full text] [doi: 10.1016/j.addbeh.2009.10.012] [Medline: 19900763]
29. Lee CM, Neighbors C, Kilmer JR, Larimer ME. A brief, web-based personalized feedback selective intervention for college student marijuana use: a randomized clinical trial. Psychol Addict Behav 2010 Jun;24(2):265-273 [FREE Full text] [doi: 10.1037/a0018859] [Medline: 20565152]
30. O'Kearney R, Gibson M, Christensen H, Griffiths KM. Effects of a cognitive-behavioural internet program on depression, vulnerability to depression and stigma in adolescent males: a school-based controlled trial. Cogn Behav Ther 2006;35(1):43-54. [doi: 10.1080/16506070500303456] [Medline: 16500776]
31. 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]
32. Bewick BM, West R, Gill J, O'May F, Mulhern B, Barkham M. Providing web-based feedback and social norms information to reduce student alcohol intake: A multisite investigation. J Med Internet Res 2010 Dec 19;12(5):e59. [doi: 10.2196/jmir.1461]
33. Bewick BM, West RM, Barkham M, Mulhern B, Marlow R, Traviss G, et al. The effectiveness of a Web-based personalized feedback and social norms alcohol intervention on United Kingdom university students: randomized controlled trial. J Med Internet Res 2013;15(7):e137 [FREE Full text] [doi: 10.2196/jmir.2581] [Medline: 23883616]
34. Fraeyman J, Van Royen P, Vriesacker B, De Mey L, Van Hal G. How is an electronic screening and brief intervention tool on alcohol use received in a student population? A qualitative and quantitative evaluation. J Med Internet Res 2012 Apr 23;14(2):e56. [doi: 10.2196/jmir.1869]
35. Mangunkusumo RT, Brug J, de Koning HJ, van der Lei J, Raat H. School-based internet-tailored fruit and vegetable education combined with brief counselling increases children's awareness of intake levels. Public Health Nutr 2007 Mar;10(3):273-279. [doi: 10.1017/S1368980007246671] [Medline: 17288625]
36. Ajzen I. The theory of planned behavior. Organizational Behavior and Human Decision Processes 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
37. Weinstein ND, Rothman AJ, Sutton SR. Stage theories of health behavior: conceptual and methodological issues. Health Psychol 1998 May;17(3):290-299. [Medline: 9619480]
38. Hawkins RP, Kreuter M, Resnicow K, Fishbein M, Dijkstra A. Understanding tailoring in communicating about health. Health Educ Res 2008 Jun;23(3):454-466 [FREE Full text] [doi: 10.1093/her/cyn004] [Medline: 18349033]
39. Oenema A, Brug J, Dijkstra A, de Weerdt I, de Vries H. Efficacy and use of an internet-delivered computer-tailored lifestyle intervention, targeting saturated fat intake, physical activity and smoking cessation: a randomized controlled trial. Ann Behav Med 2008 Apr;35(2):125-135. [doi: 10.1007/s12160-008-9023-1] [Medline: 18363076]
40. 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]
41. 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]
42. de Nooijer J, Oenema A, Kloek G, Brug J, de Vries H, de Vries NK. Bevordering van gezond gedrag via internet: Nu en in de toekomst. Maastricht, The Netherlands: Maastricht University; 2005.
43. Spence SH, Holmes JM, March S, Lipp OV. The feasibility and outcome of clinic plus internet delivery of cognitive-behavior therapy for childhood anxiety. J Consult Clin Psychol 2006 Jun;74(3):614-621. [doi: 10.1037/0022-006X.74.3.614] [Medline: 16822117]
44. 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]
45. Van Voorhees BW, Vanderplough-Booth K, Fogel J, Gladstone T, Bell C, Stuart S, et al. Integrative internet-based depression prevention for adolescents: a randomized clinical trial in primary care for vulnerability and protective factors. J Can Acad Child Adolesc Psychiatry 2008 Nov;17(4):184-196 [FREE Full text] [Medline: 19018321]
46. Hoek W, Marko M, Fogel J, Schuurmans J, Gladstone T, Bradford N, et al. Randomized controlled trial of primary care physician motivational interviewing versus brief advice to engage adolescents with an Internet-based depression prevention

intervention: 6-month outcomes and predictors of improvement. Transl Res 2011 Dec;158(6):315-325. [doi: 10.1016/j.trsl.2011.07.006] [Medline: 22061038]
47. Cuijpers P, Jonkers R, de Weerdt I, de Jong A. The effects of drug abuse prevention at school: the 'Healthy School and Drugs' project. Addiction 2002 Jan;97(1):67-73. [Medline: 11895272]
48. de Haes WFM, Schuurman JH. Resultaten van het rotterdamse drugsvoorlichtingsexperiment. Tijdschrift voor sociale geneeskunde 1975;53:394-410.
49. Horowitz JL, Garber J, Ciesla JA, Young JF, Mufson L. Prevention of depressive symptoms in adolescents: a randomized trial of cognitive-behavioral and interpersonal prevention programs. J Consult Clin Psychol 2007 Oct;75(5):693-706. [doi: 10.1037/0022-006X.75.5.693] [Medline: 17907851]
50. Statline. Voortgezet onderwijs; deelname leerlingen naar onderwijssoort. 2013. URL: http://statline.cbs.nl/StatWeb/publication/?DM=SLNL&PA=80040ned&D1=0-21&D2=0&D3=0&D4=0&D5=0&D6=0&D7=3-9&VW=T [WebCite Cache ID 6I9Oh6ofn]
51. Hanson MD, Chen E. Socioeconomic status and health behaviors in adolescence: a review of the literature. J Behav Med 2007 Jun;30(3):263-285. [doi: 10.1007/s10865-007-9098-3] [Medline: 17514418]
52. Calear AL, Christensen H. Systematic review of school-based prevention and early intervention programs for depression. J Adolesc 2010 Jun;33(3):429-438. [doi: 10.1016/j.adolescence.2009.07.004] [Medline: 19647310]
53. Borsari B, Muellerleile P. Collateral reports in the college setting: a meta-analytic integration. Alcohol Clin Exp Res 2009 May;33(5):826-838 [FREE Full text] [doi: 10.1111/j.1530-0277.2009.00902.x] [Medline: 19298324]
54. Rutter M. The development of psychopathology of depression: Issues perspectives. In: Depression in Young People: Developmental and Clinical Perspectives. New York: Guilford Press; 1986.
55. 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

CHQ-CF-GH4: Child Health Questionnaire-Child Form-General Health
RCT: randomized controlled trial
SDQ: Strengths and Difficulties Questionnaire
YSR: Youth Self Report

Edited by G Eysenbach; submitted 04.12.13; peer-reviewed by K Stasiak, A McPherson; comments to author 12.02.14; revised version received 04.03.14; accepted 30.03.14; published 30.05.14.

Please cite as:
Bannink R, Broeren S, Joosten-van Zwanenburg E, van As E, van de Looij-Jansen P, Raat H
Effectiveness of a Web-Based Tailored Intervention (E-health4Uth) and Consultation to Promote Adolescents' Health: Randomized Controlled Trial
J Med Internet Res 2014;16(5):e143
URL: http://www.jmir.org/2014/5/e143/
doi:10.2196/jmir.3163
PMID:24878521

Original Paper

Initial Outcomes From a 4-Week Follow-Up Study of the Text4baby Program in the Military Women's Population: Randomized Controlled Trial

W Douglas Evans[1], MA, PhD; Jasmine Wallace Bihm[1], MPH; Daniel Szekely[2], MD; Peter Nielsen[2], MD; Elizabeth Murray[2], MD; Lorien Abroms[1], ScD; Jeremy Snider[3], MPH

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

[2]Madigan Army Medical Center, Department of Obstetrics and Gynecology, Tacoma, WA, United States

[3]School of Public Health, University of Washington, Seattle, WA, United States

Corresponding Author:
W Douglas Evans, MA, PhD
Milken Institute School of Public Health
The George Washington University
950 New Hampshire Avenue
Washington, DC, 20052
United States
Phone: 1 202 994 3632
Fax: 1 202 994 3601
Email: wdevans@gwu.edu

Abstract

Background: The use of mobile phone technologies for health promotion and disease prevention has advanced rapidly in recent years. Text4baby is a theory-based mobile health (mHealth) program in which text messages are delivered to pregnant women and new mothers to improve their health care beliefs and behaviors and improve health status and clinical outcomes. Recent evaluations of Text4baby have found that it improves targeted health attitudes and beliefs, but effects on behavior have not yet been determined.

Objective: In this study, investigators aimed to evaluate Text4baby in the military women's population.

Methods: Investigators conducted a randomized controlled trial at Madigan Army Medical Center in Tacoma, Washington, from December 2011 through September 2013. All participants were pregnant women first presenting for care at Madigan. Investigators conducted a baseline assessment using a 24-item, self-administered online survey of attitudes and behaviors related to Text4baby message content. Participants were randomized to Text4baby plus usual care (intervention) or usual care alone (control). Investigators analyzed treatment effects of Text4baby on short-term targeted outcomes 4 weeks post enrollment.

Results: For this study, 943 patients were randomized and completed a baseline assessment. The average patient age was 28 years and nearly 70% self-identified as Caucasian. 48.7% of enrollees (459/943) completed the first follow-up assessment. Higher rates of single and working/in-school patients dropped out of the intervention arm of the study, and we adjusted for this finding in subsequent models. However, while investigators were unable to re-survey these participants, only 1.9% of Text4baby enrollees (18/943) dropped the service during the study period. Adjusted and unadjusted logistic generalized estimating equation models were developed to assess intervention effects on measured outcomes. In the model adjusting for age, marital status, having had a previous baby, and race/ethnicity, there was a significant effect of Text4baby intervention exposure on increased agreement with belief in the importance of taking prenatal vitamins (OR 1.91, 95% CI 1.08-3.34, *P*=.024). All of these attitudes had been targeted by at least one text message during the 4-week evaluation period examined in this study. In unadjusted models, there was a significant effect of intervention exposure on belief in the importance of visiting a health care provider to be a healthy new mother (OR 1.52, 95% CI 1.01-2.31, *P*=.046) and in the health risks of alcohol during pregnancy (OR 2.06, 95% CI 1.00-4.31, *P*=.05). No behavioral effects of the intervention were observed in this analysis.

Conclusions: Text4baby is a promising program that offers lessons for future mHealth activities. This large-scale study demonstrated initial effects of the program on attitudes and beliefs targeted by the messages received by women during the study period. Results confirm previous findings from Text4baby studies and other mHealth research. Future analyses will examine dosage effects of the intervention on behaviors and clinical outcomes.

(J Med Internet Res 2014;16(5):e131) doi:10.2196/jmir.3297

KEYWORDS

Text4baby; prenatal health care; mobile health; military health; health behavior

Introduction

In recent years, the use of mobile phone technology for health and health care (or mHealth) has grown rapidly as a strategy to increase patient adherence to care protocols for chronic conditions. Mobile phones allow patients to monitor and take control of their own health. A number of research studies have been conducted to test the outcomes of mobile-enabled health behavior change, drug and treatment adherence, and health care solutions, which have generally indicated that mobile solutions are effective in promoting health behaviors and health care treatment and adherence outcomes [1,2].

Despite the rapid growth of mHealth, there is relatively little research on its role as a channel through which to deliver health promotion interventions. There is also little research on the effects of mHealth exposure and dosage on outcomes [3]. Some mHealth programs have been developed as reminder systems, such as in disease management and treatment adherence programs in smoking cessation, weight loss, and diabetes self-management [4-7]. In some limited cases, mHealth approaches such as text messaging have been shown to be effective in promoting adoption of behaviors such as exercise, nutrition, family planning, and encouraging risk behavior avoidance, such as smoking cessation [6-10].

Texting, a relatively simple approach to mHealth, has been the most widely used and evaluated form of mHealth intervention [11]. Text messages can act as motivators, cues to action, and also reminder mechanisms to reinforce messages [12]. The latter use, as a reminder system to reinforce treatment adherence, is perhaps the most widely used application of text messaging as an mHealth intervention [13,14].

In the past, relatively few studies have been conducted to test the efficacy of mHealth programs as a channel for full-fledged health promotion programs. However, this situation has begun to change, one sign of which is heightened interest in mHealth theory and development of behavior change models that take the mobile phone as a delivery channel into account [5,15]. Mobile phones have unique characteristics as a channel—their ubiquity, mobility, constant availability, and multiple media modalities, among others [2,15]. The future of mHealth research is to build an evidence base that supports extensions and advancement of existing behavioral theory to address how people use mobile devices and the effects of the channel on health outcomes.

The Text4baby mHealth program [16] is an example of using text messaging for health promotion and behavior change. The service launched in February 2010 and delivers text messages to pregnant women and new mothers. The program consists of 135 distinct prenatal text messages (and a separate library of postpartum messages for new mothers, not addressed here) delivered on a schedule timed to mothers' enrollment and due date of the baby. In this way, messages are tied to the information most needed during a particular stage of pregnancy. The program is directed by the National Healthy Mothers, Healthy Babies Coalition, and text messages and participant data are managed by Voxiva, Inc. It represents one of the largest mHealth text-based programs developed to date. Several recent studies of the Text4Baby program have found that Text4Baby changes attitudes and beliefs, but effects on behavior have not yet been established [10,17,18].

Text4baby also offers an example of the development of behavioral theory and its adaptation to the design mHealth interventions. First, following the health belief model (HBM), Text4baby messages serve as a cue to prenatal care action and behavior change targeting primarily pregnant women [19]. Second, text messaging can spread social influence and diffusion of information within a target population [20]. In Text4baby, messages diffuse risk information, such as the dangers of smoking and secondhand smoke during pregnancy and serve to reinforce other social influences, such as health care providers, friends, and family. Third, Text4baby uses social cognitive theory (SCT) to build participants' self-efficacy and provide social models of behavior through text messaging, which develop a Text4baby brand identity and brand equity (eg, the brand's market position, consumer loyalty) [21]. An mHealth program can serve to create a brand identity as a "friend" and "trusted source" that young women come to rely on. Messages are authentic, fun, and promote ways to live a satisfying lifestyle while reducing risk and promoting maternal and child health [17]. Audiences must develop a connection with messages and find them to be credible sources that they can rely on in order to be successful [22-24]. Future mHealth programs can build on these lessons to develop mobile health messages that address the quality of life and other motivational factors to help make recommended health behaviors fun, easy, and popular with audiences.

Over the past few years, the military has made major investments in mHealth programs to promote health and improve treatment among enlisted personnel and their families [25]. The target population, which includes active duty military personnel, spouses, and family members, face significant additional stressors beyond those in the civilian population that may affect maternal and child health outcomes. Female service members struggle to balance family and service and are much more likely to be a single parent than male troops [26]. It is estimated that about 10% of female military personnel become pregnant each year, and approximately 75,000 military children are estimated to be younger than 1 year of age [26]. Thus, there is a large Military Health Service (MHS) population that could benefit from the Text4baby program, as it is a free resource that can help support military families and address a growing demand for maternal and child health and health care services.

The current study examines the first phase of a randomized controlled trial (RCT) to evaluate Text4baby in the military women's population at Madigan Army Medical Center

(Madigan) in Tacoma, Washington. The first specific aim of this study was to describe the Madigan Text4baby RCT, its design, methods, and baseline (BL) data. The second specific aim was to evaluate initial outcomes of the program at 4 weeks post BL. This is also to confirm a previous smaller-scale evaluation of Text4baby conducted by the investigators in a largely Spanish-speaking immigrant population in Virginia [10].

Regarding the second aim, there were two specific hypotheses. First (H1), Text4baby will be effective in improving prenatal knowledge, attitudes, and beliefs (KAB) targeted by messages delivered during the study period. Because this was a study of short-term Text4baby effects, analysis focused on Text4baby effects on KAB targeted during the initial study period, such as taking prenatal vitamins, increased fruit and vegetable intake, visiting health care providers, and avoiding smoking or drinking. Second (H2), Text4baby will be effective in improving prenatal behaviors targeted by messages.

Methods

Design and Measures

The investigators conducted an RCT of Text4baby prenatal messages at Madigan Army Medical Center (Madigan), a large tertiary-care Army Medical Center in Tacoma, Washington, serving Joint Base Lewis-McCord. Female military health care beneficiaries age 18-45 years (both active duty and family members) who first presented for prenatal care at Madigan prior to 14-weeks gestation were eligible for the study. The date of pregnancy was established by last menstrual period (LMP) dating. Additional inclusion criteria included having a working cell phone and speaking and reading fluent English. Female patients meeting the inclusion/exclusion criteria were recruited for the study at the end of their initial prenatal care visit to the Madigan Obstetrics and Gynecology clinic. Following medical consultation, the health care provider asked if the patient would be willing to participate in the research study. Those who agreed met with a member of the research team in a private space in the clinic, underwent a written informed consent, and then were offered the opportunity to complete a BL survey on a password-protected computer in a private room in the clinic. Three dedicated laptops were available for participants to complete the online survey. No compensation was provided for participation.

Individuals who needed more time to decide whether to participate were allowed to leave and provide a final decision by calling or texting back. A card was provided with the name and number of Text4baby points of contact they could reach to discuss the study. Those who agreed to participate after leaving the clinic completed the BL survey online at a secure, password-protected website.

After BL survey completion, participants were assigned to a study condition (treatment or control). Using an algorithm that generated a randomized list of individual assignments to treatment or control condition for a sample of 996, investigators assigned participants either to continue with usual prenatal care (control) or receive usual care plus enroll in Text4baby (treatment). After BL, participants in both groups were surveyed again after 4 weeks. Subsequent follow-ups were conducted but are not reported here. BL data collection started in December 2011 and ended in January 2013. Follow-up data collection for the 4-week follow-up survey began in January 2012 and ended in March 2013. This study reports on these BL and 4-week follow-up data.

Participants had the option to take each of the follow-up surveys in the clinic after subsequent health care appointments or remotely, online using their own computer/device. All surveys were completed anonymously and linked by an assigned participant ID. Clinicians who met with patients were blinded—the randomization occurred outside the actual clinical visit and the trial data were not accessed by the clinicians during the study. The onsite trial coordinator responsible for data collection used the random assignment list to direct participants in the Text4baby arm to enroll in the service and was aware of participant treatment or control status.

The BL survey instrument was a 24-item Web-based tool developed by the research team that contained a battery of items drawn from previous research [10]. Variables for behavioral outcomes were derived from validated instruments, including the Behavioral Risk Factor Surveillance Survey (BRFSS) and National Health and Nutritional Examination Survey (NHANES).

The survey instrument contained a series of questions on participant attitudes, beliefs, and behaviors related to the text messages contained in Text4baby*:* nutrition, smoking, taking vitamins, alcohol use, flu shots, health care appointments, health information seeking, and related risk prevention behaviors. The instrument also captured confirmed recall, reactions, and receptivity to the text messages based on validated measures previously published [27,28]. Additionally, demographic information such as age, race, ethnicity, sponsor rank, marital status, and parity (previous live birth by the mother) was collected from the medical record. The study protocol was approved as minimal risk research by the Madigan Institutional Review Board on July 26, 2011, and served as the Institutional Review Board (IRB) of record for the George Washington University, which entered into an institutional agreement with the Department of Defense (DOD) Human Research Protection Program on August 1, 2011.

Text4baby Intervention

During enrollment in the clinic, participants assigned to the control condition were excused immediately after completing the BL survey. Those assigned to the treatment condition were immediately directed to enroll in Text4baby by texting a phrase that tagged them as participants in the Madigan study to a designated SMS (short message service) short code to receive messages for the duration of their pregnancy (or until they dropped out of the program). This combination of enrollment phrase and short code had been established in the Voxiva Text4baby data system to identify participants as members of the Madigan evaluation study. This system ensured that only Text4baby participants who were enrolled in the Madigan study were counted in our treatment group. We monitored the control group through the Voxiva data system to ensure that none of

these participants separately enrolled for Text4baby outside of our study, and none did.

The Voxiva system consists of data obtained from participants during their enrollment: mobile phone number, due date (in order to time the pregnancy text messages appropriately to the gestational age of the baby), and participant's home zip code [29]. Voxiva did not have access to any data collected through this evaluation or any patient information stored at Madigan. Once enrolled, Text4baby participants received welcome and introductory messages and then began receiving the three text messages per week throughout their enrollment. These messages were tailored to the date of enrollment and baby's gestational age. Thus a woman enrolling in her 10th week of pregnancy would begin receiving week 10 messages.

If at any time during the pregnancy, a fetal loss occurred, patients were disenrolled from Text4baby and appropriate perinatal grief counseling was offered to the patient and her partner. Patients were also provided the option to disenroll from Text4baby by texting "STOP" from their cell phones if they no longer wanted to participate in the program.

Sampling

The sampling frame consisted of all pregnant women who were military health care beneficiaries first presenting for initial prenatal care at Madigan (ie, had not visited another health facility for the current pregnancy). Participants were further required to be within the first 14 weeks of their baby's gestational age. All others were excluded from the study. We drew a random sample of all women meeting criteria who first presented for care at the Madigan Obstetrics and Gynecology Clinic between December 1, 2011, and January 31, 2013. Recruitment from the sample occurred on a rolling basis, meaning that they were recruited over time until the targeted sample (as determined by statistical power requirements) was reached. Previous interventions to promote reproductive health care utilization among low- and middle-income women suggest an approximate 12% effect (intervention versus control) of such programs after a 12-month time period [30]. Power analysis estimated the required sample to be 996 participants in total; 249 participants per study condition and per stratum, assuming a 10% attrition rate at 1-year follow-up.

Data Collection Procedures

In November 2011, before the start of data collection, the research team held an introductory meeting and training session with clinical staff at Madigan to go over the study purpose and protocol. BL surveys were generally completed in clinic, with a small number completed remotely, as previously described. Follow-up (FL) surveys were primarily completed remotely online. In an attempt to ensure high response and long-term retention, we used several methods to address noncontact and noncooperation, including (1) text messages for FL data collection notification, (2) a local phone number for participants if there were questions or a desire to schedule a time to speak with the investigators (provided a business card to participants during enrollment), (3) assurances of confidentiality, and (4) a nurse available to answer the phone and take the messages to convey to investigators. One week prior to each FL survey, the onsite research team member sent text messages reminding participants of the upcoming survey FL date and offered them the option to take the survey remotely if they would not be at the clinic. Participants were considered to have quit the study if they were unreachable after several attempts, and no further FL attempts were made. At follow-ups, the same behavioral survey instrument was repeated with the addition of a short battery of questions about exposure and reactions to the Text4baby campaign and its messages.

Data Analysis

Stata Version 12 was used in all analyses. Descriptive statistics were calculated for all outcomes and demographic variables. Cross-tabulations of these same variables by study condition and survey time point were also calculated.

Generalized estimating equations (GEE) logistic regression was used to construct separate models for each of the attitudes, beliefs, and behavioral outcomes. Investigators estimated the odds of positive change over time in response to each of the behavioral outcome variables as a function of Text4baby text message exposure through use of an interaction term including program exposure and progression to follow-up measurement.

In addition to an unadjusted model, which strictly looked at the effect of the randomized intervention on those who completed both BL and follow-up interviews (n=459), a second adjusted model included several maternal covariates: age quintile, parity, marital status, and race. For missing data and attrition of participants, a *t* test was used to compare covariates, including sociodemographic and other variables used in the regressions, between cases with and without missing data to verify whether or not data were missing completely at random. It was determined that both maternal race and marital status, which were both missing for 22.7% (215/943) of BL participants, were potentially variables that were differentially missing for women of certain racial and marital statuses. Therefore, a multiple imputation (MI) model was constructed to account for missing race and marital status, through use of a logit function, with parity, age, and treatment status as predictors of both race and marital status [31].

The MI model also accounted for attrition through inverse probability weighting, by predicting likelihood of dropout through use of a logistical regression model, using parity, age, and imputed or actual race and marital status as predictors [32]. This regression model was then used to predict the probability of a missing case and to assign higher weights to complete cases of individuals in the sample who represent those more likely to have dropped out of the study. Sixty repetitions of the model were implemented, and results from each imputation averaged to produce point estimates for an intervention effect.

Results

Of 1078 women who entered the Madigan OB/Gyn clinic during the study period, 996 were asked to participate, or 92.39%. Of these, 94.7% completed a BL survey online (943/996). Among the BL participants, 48.7% completed a 4-week follow-up survey (459/943).

Table 1 provides the BL sample characteristics. Overall, the sample was predominantly Caucasian (69.6%, 656/943), with an average age of 26.53 years. A majority (63.1%, 595/943) reported currently attending school or working outside of the home. Over 70% (663/943) of the participants reported being married. Close to half of the participants (47.8%, 451/943) reported having had a baby prior to this pregnancy. In terms of sponsor rank, the vast majority of participants were enlisted service members or a dependent family member of an enlisted service member (86.8%, 819/943) and the remainder commissioned or warrant officers (13.2%, 124/943).

Equivalence of means at baseline was tested by comparing the baseline treatment and control condition samples. The comparison revealed a larger, statistically significant percent reporting smoking in the last 30 days: 15.34%, (95% CI 12.08-18.58) in the control versus 9.64% (95% CI 6.95-12.32) in the treatment group, respectively. There was also a larger, statistically significant percent who reported consuming 3 or more vegetables per day in the control versus treatment group: 37.82% (95% CI 33.44-42.19) in control versus 29.98% (95% CI 25.81-34.15] in treatment, respectively.

Because of the significant loss-to-follow-up, demographic characteristics of the BL sample were compared to those who remained in the study at follow-up. That analysis revealed statistically significant differences in those reporting being married (70.31% at BL vs 76.69% at FL, P=.000) and in those currently working or attending school (63.10% at BL vs 53.16% at FL, P<.001). No differential attrition between treatment and control conditions was observed.

Table 2 presents cross-tabulations of the outcome variables by study condition and survey time point. There were significant improvements in several outcome attitudes, beliefs, and behaviors from BL to FL. However, no significant differences in improvements over time were observed between the treatment and control study conditions. Increases in strong agreement with beliefs in the importance of prenatal vitamins and health risks of drinking alcohol were observed between BL and FL. An increase in self-reported searching for health information online and consuming 3 or more serving of fruits and vegetables per day was observed. Finally, a decrease in self-reported cigarette smoking was observed between the two time points.

Table 3 summarizes the results of the GEE logistic model for the intervention effects BL to FL on measured outcomes. There was a significant treatment effect for improvement in strong agreement with the statement that “If I visit my health care provider on a regular basis, I will be a healthy new mother” in the complete-case, unadjusted model, with a greater likelihood of improvement to strong agreement in the intervention group (OR 1.52, 95% CI 1.01-2.31, P=.046). There was also a significant treatment effect for improvement in strong agreement with the statement that “drinking alcohol will harm the health of my developing baby” in the unadjusted model, with a greater likelihood of improvement to strong agreement in the intervention group (OR 2.06, 95% CI 1.00-4.31, P=.05). There were marginally significant effects, at the P=.10 level, for improvement in strong agreement for beliefs about the importance of eating fruits and vegetables, and in taking prenatal vitamins, to the health of the developing baby.

In the adjusted model, which accounts for four socioeconomic variables, imputations for missing values for marital status and race, and inverse probability weighting to account for the noted attrition, there were marginally significant effects for improved strong agreement with the statement “If I visit my health care provider on a regular basis, I will be a healthy new mother” as well as “Drinking alcohol will harm the health of my developing baby”. There was also a significant treatment effect for improvement in strong agreement with the statement “Taking prenatal vitamins will improve the health of my developing baby”, with a greater likelihood of improvement to strong agreement in the intervention group (OR 1.91, 95% CI 1.08-3.34, P=.024). There were no effects of the Text4baby intervention on any of the measured behaviors at FL.

Table 1. Baseline sample descriptive statistics (N=943).

Variables	n	%
Age, years	943	26.5
<20	31	3.3
20-34	837	88.8
35+	75	7.9
Race		
White	656	69.6
Black	75	7.9
Asian-Pacific Islander	25	2.6
Western Hemisphere Indians	2	0.2
Other/Unknown	185	19.6
Ethnicity		
Filipino	206	21.8
Hispanic	53	5.6
Other Asian/Pacific Islander	19	2.0
Southeast Asian	6	0.6
Other/Unknown/Not Hispanic	659	69.9
Marital status		
Single/Never married	72	7.6
Married	663	70.3
Separated/Divorced/Widowed	7	0.7
Unknown/Null	201	21.3
Sponsor rank		
Enlisted	819	86.8
Commissioned Officers	107	11.3
Warrant Officers	13	1.4
Null	4	0.4
Parity		
No	492	52.2
Yes	451	47.8
Pre-pregnancy Body Mass Index	327	27.2
Underweight	7	0.7
Normal	154	16.4
Overweight	97	10.3
Obese	63	6.7
Ever participated in WIC Program[a]	320	33.9
Currently in school or working outside the home	595	63.1
Ever gone online to search for prenatal care information	711	75.4

[a]WIC=Nutritional Program for Women, Infants, and Children.

Table 2. Bivariate pre-post comparison of measured outcome variables by treatment group.

	Baseline sample (n=459)				Follow-up sample (n=459)				
Attitudes, Strongly Agree	Mean, %	95% CI	Control (n=230), %	Text4baby (n=229), %	Mean, %	95% CI	Control (n=230), %	Text4baby (n=229), %	*P* value[a]
Eating 5 or more fruits and vegetables per day is important to the health of my developing baby	68.41	63.94-72.64	70.43	66.38	67.10	62.60-71.39	64.78	69.43	.672
Taking a prenatal vitamin is important to the health of my developing baby	88.02	84.69-90.84	89.57	86.46	78.00	73.92-81.70	76.09	79.91	.000
I am prepared to be a new mother	52.94	48.26-57.58	55.22	50.66	52.07	47.39-56.72	53.49	50.66	.792
If I visit my health care provider on a regular basis, I will be a healthy new mother	48.37	43.71-53.04	50.00	46.72	44.44	39.84-49.12	40.87	48.03	.234
If I visit my health care provider on a regular basis, my baby will be healthy	44.88	40.27-49.56	43.48	46.29	42.92	38.34-47.59	39.13	46.72	.550
Smoking will harm the health of my developing baby	92.59	89.80-94.82	92.61	92.58	89.98	86.86-92.57	87.83	92.14	.161
Secondhand smoke will not harm the health of my developing baby.	18.08	14.67-21.91	18.70	17.47	18.95	15.47-22.84	18.70	19.21	.734
Drinking alcohol will harm the health of my developing baby	90.41	87.35-92.95	93.48	87.33	85.62	82.07-88.70	85.65	85.59	.025
Taking a prenatal vitamin will improve the health of my developing baby	76.25	72.09-80.07	74.78	77.73	69.50	65.06-73.68	64.78	74.24	.021
Behaviors									
Have you ever gone online to search for prenatal care information?	10.02	7.43-13.14	9.57	10.48	16.78	13.47-20.51	15.21	18.34	.003
In last 30 days, did you smoke?	11.98	9.16-15.31	15.65	8.30	4.36	2.68-6.65	6.09	2.62	.000
Since you found out about your pregnancy, have you consumed alcoholic beverages?	2.40	1.20-4.25	1.74	3.06	2.61	1.36-4.52	2.61	2.62	.833
Ate 3 or more servings of fruit a day	35.95	31.55-40.53	37.39	34.50	43.36	38.77-48.03	43.48	43.23	.022
Ate 3 or more servings of vegetables a day	32.24	27.99-36.73	33.48	31.00	33.12	28.82-37.63	30.87	35.37	.779

[a]*P* value presented represents the difference between the baseline and follow-up sample mean.

Table 3. Effects of Text4baby and covariates on improvements in outcome variables from baseline to follow-up.

	Effect of intervention and time on strong agreement (unadjusted)	Effect of intervention and time on strong agreement (fully adjusted, for age [quintile], parity, imputed marital status and race, and use of inverse probability weighting to account for attrition)[a]
	(n=459)	(n=943 for baseline sample)
	OR (95% CI), *P* value	OR (95% CI), *P* value
Attitudes		
Eating 5 or more fruits and vegetables per day is important to the health of my developing baby	1.49 (0.96-2.31), *P*=.075	1.47 (0.83- 2.63), *P*=.189
Taking a prenatal vitamin is important to the health of my developing baby	1.68 (0.96-2.94), *P*=.069	1.73 (0.80-3.73), *P*=.164
I am prepared to be a new mother	1.07 (0.673-1.57), *P*=.804	1.28 (0.74-2.23), *P*=.555
If I visit my health care provider on a regular basis, I will be a healthy new mother	1.52 (1.01-2.31), *P*=.046	1.66 (0.98-2.81), *P*=.058
If I visit my health care provider on a regular basis, my baby will be healthy	1.22 (0.83-1.80), *P*=.320	1.32 (0.81-2.16), *P*=.268
Smoking will harm the health of my developing baby	1.63 (0.74-3.61), *P*=.226	2.25 (0.64-7.92), *P*=.204
Secondhand smoke will not harm the health of my developing baby (reverse coded*)	1.14 (0.81-1.58), *P*=.450	0.82 (0.47-1.44), *P*=.491
Drinking alcohol will harm the health of my developing baby	2.06 (1.00-4.31), *P*=.050	2.19 (0.87-5.52), *P*=.095
Taking prenatal vitamins will improve the health of my developing baby	1.33 (0.84-2.10), *P*=.221	1.91 (1.08-3.34), *P*=.024
Behaviors		
Have you ever gone online to search for prenatal care information?	0.91 (0.45-1.87), *P*=.810	0.72 (0.21-2.43), *P*=.592
In the last 30 days, did you smoke?	0.86 (0.38-1.97), *P*=.726	1.16 (0.41-3.27), *P*=.768
Since you found out about your pregnancy, have you consumed alcoholic beverages?	0.49 (0.15-1.61), *P*=.241	0.57 (0.14-2.36), *P*=.436
Ate 3 or more servings of fruit a day	1.04 (0.68-1.60), *P*=.842	1.07 (0.63-1.84), *P*=.758
Ate 3 or more servings of vegetables a day	1.29 (0.83-1.97), *P*=.249	1.17 (0.42-3.27), *P*=.795

[a]Indicates change to “strongly disagree” (reverse coded).

Discussion

Principal Findings

Text4baby is now a widely known mHealth program that has demonstrated the rapid scalability of behavior change interventions using the mobile phone. Since its launch in February 2010, over 670,000 individuals have enrolled in the service as of January 2014 [33]. This widespread adoption suggests that the program has broad appeal and may represent a valuable health promotion model in the area of maternal and child health [34]. Given widespread adoption, it is important to evaluate the effectiveness of such programs in changing health behaviors and affecting other health outcomes.

This pilot study examined short-term effects of Text4baby exposure 4 weeks post enrollment on attitudes, beliefs, and behaviors targeted by the text messages. Overall, we found that exposure to the text messages improved some targeted attitudes and beliefs, which partially confirms H1. Specific targeted beliefs, including those about the importance of prenatal health care, the risks of alcohol use during pregnancy, and the importance of prenatal vitamins were more likely to improve given exposure to Text4baby. These results were consistent with findings from a previous randomized controlled evaluation of Text4baby [10]. Each of these beliefs was targeted by some of the text messages delivered during the intervention period.

Thus there is substantial evidence of cognitive changes associated with Text4baby. Following the program’s theory of change [10], these changes would be hypothesized to mediate subsequent behavior changes. However, in this short-term follow-up assessment, no behavior changes were observed in this study. These results disconfirm H2; no behavioral changes

were observed for any of the behavioral outcomes targeted by the text messages.

Clearly behavior change is the objective of mobile health promotion programs such as Text4baby. However, behavioral effects often require longer intervention periods in order to manifest [35,36]. While Text4baby was not explicitly designed on the transtheoretical model (TTM), that theory would suggest that pregnant women may be rapidly progressing through pre-contemplation to contemplation to behavioral initiation, but somewhat longer periods of time than the 4-week FL period examined in this study may be required for behavior change. As the pregnancy would generally continue for several more months, the action phase according to TTM would be more likely to occur at subsequent time points. Future studies based on longer evaluation periods may shed light on potential behavioral effects of the Text4baby program.

The implications of this study for Text4baby are further confirmation of the program's conceptual model, published elsewhere, that is based on HBM and SCT [10]. Thus this evaluation is one step toward validating a new theoretical approach to mHealth programs—one that calls for additional research and theoretical investigation in the field. Previous communication research suggests that targeted health communications delivered using validating messaging strategies may, by themselves, have small but statistically significant effects on subsequent health cognitions and behavior [37,38]. The theory behind Text4baby, then, is that beliefs targeted by the program's text messages will have beneficial effects on those specific beliefs, which in turn will be associated with improvements related to health behaviors. This study provides confirmation of the first association but leaves open whether the program can demonstrate behavioral effects. Future studies should also examine whether targeted beliefs that improve as a result of Text4baby exposure mediate effects on behavioral outcomes.

Limitations

This study had two important limitations. First, there were observed differences in the BL versus follow-up samples due to attrition, with more unmarried and fewer employed or in school retained at follow-up, though not between study conditions. Single mothers and those working or in school (among the individuals who would potentially most benefit fromText4baby*)* were least likely to be retained. This may have reduced observed effects of the intervention.

Second there was higher than expected attrition from the study, at just over 50%. While very few women appear to have dropped out of Text4baby messaging, we were unable to re-contact many women. As a result, the analysis included weighting and MI procedures to account for missing data in the analyses. As a result of the attrition, statistical power was reduced as well as the ability to detect potential significant differences over time between conditions resulting in relatively wide confidence intervals for observed significant results. This limitation should be considered in light of the study's purpose as a pilot evaluation.

Conclusions

This study found that Text4baby participation improved attitudes and beliefs targeted by the intervention—specifically beliefs in the importance of prenatal vitamins in the adjusted model. There were also improvements in beliefs about attending health care appointments and about the importance of avoiding alcohol during pregnancy in the unadjusted model. Despite these improvements in beliefs targeted by Text4baby, there were no short-term effects on self-reported behavior. Future studies based on this RCT will examine whether long-term participation in Text4baby affects behaviors, and whether there is a dose-response relationship (ie, receiving more messages and/or participating for a longer period is related to outcomes). Future studies should also examine the potential for mediation effects (ie, whether targeted attitudes and beliefs mediate the relationships between exposure to Text4baby and behavioral outcomes). Finally, future studies will examine whether Text4baby participation has an effect on clinical maternal and child health outcomes.

Acknowledgments

This study was supported by grant #W81XWH-10-2-0142 to The George Washington University from the Telemedicine and Advanced Technology Research Center (TATRC), US Army Medical and Materiel Command. The authors gratefully acknowledge the valuable review and comments on this study provided by Ms Cynthia Barrigan of TATRC. The views expressed are those of the authors and do not reflect the official policy of the Department of the Army, the Department of Defense, or the US Government.

Authors' Contributions

W Douglas Evans was the principal investigator of the study, conceptualized the manuscript, directed the analysis, and wrote sections of the main narrative. Jasmine Wallace Bihm managed day-to-day study operations, assisted in analysis, and wrote sections of the main narrative. Daniel Szekely and Peter Nielsen supervised the study at Madigan and edited sections of the main narrative. Jeremy Snider conducted multivariable analysis and wrote sections of the main narrative. Lorien Abroms edited sections of the main narrative.

Conflicts of Interest

None declared.

References

1. Cole-Lewis H, Kershaw T. Text messaging as a tool for behavior change in disease prevention and management. Epidemiol Rev 2010 Apr;32(1):56-69 [FREE Full text] [doi: 10.1093/epirev/mxq004] [Medline: 20354039]
2. Abroms LC, Ahuja M, Kodl Y, Thaweethai L, Sims J, Winickoff JP, et al. Text2Quit: results from a pilot test of a personalized, interactive mobile health smoking cessation program. J Health Commun 2012;17 Suppl 1:44-53 [FREE Full text] [doi: 10.1080/10810730.2011.649159] [Medline: 22548598]
3. Evans WD. Branding Social and Health Behavior: An Education and Research Agenda. In: Evans WD, editor. Psychology of Branding. Hauppage, NY: Nova Science Publishers; 2013:133-148.
4. Yoon KH, Kim HS. A short message service by cellular phone in type 2 diabetic patients for 12 months. Diabetes Res Clin Pract 2008 Feb;79(2):256-261. [doi: 10.1016/j.diabres.2007.09.007] [Medline: 17988756]
5. Kim SI, Kim HS. Effectiveness of mobile and internet intervention in patients with obese type 2 diabetes. Int J Med Inform 2008 Jun;77(6):399-404. [doi: 10.1016/j.ijmedinf.2007.07.006] [Medline: 17881285]
6. Newton RU, Taaffe DR, Spry N, Gardiner RA, Levin G, Wall B, et al. A phase III clinical trial of exercise modalities on treatment side-effects in men receiving therapy for prostate cancer. BMC Cancer 2009;9:210 [FREE Full text] [doi: 10.1186/1471-2407-9-210] [Medline: 19563641]
7. Patrick K, Raab F, Adams MA, Dillon L, Zabinski M, Rock CL, et al. A text message-based intervention for weight loss: randomized controlled trial. J Med Internet Res 2009;11(1):e1 [FREE Full text] [doi: 10.2196/jmir.1100] [Medline: 19141433]
8. Mackenzie H. A text messaging trial in family planning clinics. Stud Health Technol Inform 2009;146:154-159. [Medline: 19592826]
9. Abroms LC, Windsor RA, Simons-Morton B. Getting young adults to quit smoking: a formative evaluation of the X-Pack Program. Nicotine Tob Res 2008 Jan;10(1):27-33 [FREE Full text] [doi: 10.1080/14622200701767852] [Medline: 18188742]
10. Evans WD, Wallace JL, Snider J. Pilot evaluation of the text4baby mobile health program. BMC Public Health 2012;12:1031 [FREE Full text] [doi: 10.1186/1471-2458-12-1031] [Medline: 23181985]
11. Lunny C, Taylor D, Memetovic J, Wärje O, Lester R, Wong T, et al. Short message service (SMS) interventions for the prevention and treatment of sexually transmitted infections: a systematic review protocol. Syst Rev 2014;3:7 [FREE Full text] [doi: 10.1186/2046-4053-3-7] [Medline: 24433348]
12. Neville R, Greene A, McLeod J, Tracey A, Tracy A, Surie J. Mobile phone text messaging can help young people manage asthma. BMJ 2002 Sep 14;325(7364):600 [FREE Full text] [Medline: 12228151]
13. Krishna S, Boren SA. Diabetes self-management care via cell phone: a systematic review. J Diabetes Sci Technol 2008 May;2(3):509-517 [FREE Full text] [Medline: 19885219]
14. Abroms LC, Lee Westmaas J, Bontemps-Jones J, Ramani R, Mellerson J. A content analysis of popular smartphone apps for smoking cessation. Am J Prev Med 2013 Dec;45(6):732-736. [doi: 10.1016/j.amepre.2013.07.008] [Medline: 24237915]
15. Riley WT, Rivera DE, Atienza AA, Nilsen W, Allison SM, Mermelstein R. Health behavior models in the age of mobile interventions: are our theories up to the task? Transl Behav Med 2011 Mar;1(1):53-71 [FREE Full text] [doi: 10.1007/s13142-011-0021-7] [Medline: 21796270]
16. Txt4baby. URL: https://www.text4baby.org/ [accessed 2014-05-15] [WebCite Cache ID 6PakRbX7o]
17. Parker RM, Dmitrieva E, Frolov S, Gazmararian JA. Text4baby in the United States and Russia: an opportunity for understanding how mHealth affects maternal and child health. J Health Commun 2012;17 Suppl 1:30-36. [doi: 10.1080/10810730.2011.649162] [Medline: 22548596]
18. Gazmararian JA, Yang B, Elon L, Graham M, Parker R. Successful enrollment in Text4Baby more likely with higher health literacy. J Health Commun 2012;17 Suppl 3:303-311. [doi: 10.1080/10810730.2012.712618] [Medline: 23030578]
19. Glanz K, Rimer BK, Viswanath K. Health behavior and health education: theory, research, and practice. San Francisco: Jossey-Bass; 2008.
20. Rogers EM. Diffusion of innovations. New York: Free Press; 2003.
21. Bandura A. Health promotion by social cognitive means. Health Educ Behav 2004 Apr;31(2):143-164. [doi: 10.1177/1090198104263660] [Medline: 15090118]
22. Hastings G. Public health branding: applying marketing for social change. Oxford: Oxford University Press; 2008.
23. Evans WD, Davis KC, Ashley OS, Blitstein J, Koo HP, Zhang Y. Efficacy of abstinence promotion media messages: findings from an online randomized trial. J Adolesc Health 2009 Oct;45(4):409-416. [doi: 10.1016/j.jadohealth.2009.02.014] [Medline: 19766947]
24. Evans WD, Uhrig J, Davis KC, McCormack L. Efficacy methods to evaluate health communication and marketing campaigns. J Health Commun 2009 Jun;14(4):315-330. [doi: 10.1080/10810730902872234] [Medline: 19466645]
25. Poropatich R, Lai E, McVeigh F, Bashshur R. The U.S. Army Telemedicine and m-Health Program: making a difference at home and abroad. Telemed J E Health 2013 May;19(5):380-386. [doi: 10.1089/tmj.2012.0297] [Medline: 23537383]
26. Mulhall E. Iraq Women's Veterans of America (IAVA). Washington, DC: US Department of Defense Issue Report; 2009. Women Warriors: Supporting She Who has Borne the Battle URL: http://media.iava.org/IAVA_WomensReport_2009.pdf [accessed 2014-05-15] [WebCite Cache ID 6Pag4pmhy]

27. Southwell BG, Barmada CH, Hornik RC, Maklan DM. Can we measure encoded exposure? Validation evidence from a national campaign. J Health Commun 2002;7(5):445-453. [doi: 10.1080/10810730290001800] [Medline: 12455763]
28. Parker R, Ratzan SC. Health literacy: a second decade of distinction for Americans. J Health Commun 2010;15 Suppl 2:20-33. [doi: 10.1080/10810730.2010.501094] [Medline: 20845190]
29. Jordan ET, Ray EM, Johnson P, Evans WD. Early Results: Text4baby Program Reaches the Intended Audience. Nursing for Women's Health 2011;15(3):206-212.
30. Barnet B, Liu J, DeVoe M, Duggan AK, Gold MA, Pecukonis E. Motivational intervention to reduce rapid subsequent births to adolescent mothers: a community-based randomized trial. Ann Fam Med 2009;7(5):436-445 [FREE Full text] [doi: 10.1370/afm.1014] [Medline: 19752472]
31. Royston P. Multiple imputation of missing values. Stata Journal 2004;4(3):227-241.
32. Seaman SR, White IR, Copas AJ, Li L. Combining multiple imputation and inverse-probability weighting. Biometrics 2012 Mar;68(1):129-137 [FREE Full text] [doi: 10.1111/j.1541-0420.2011.01666.x] [Medline: 22050039]
33. National Healthy Mothers, Healthy Babies Coalition. Recent news from text4baby: Get involved. 2014. URL: https://www.text4baby.org/index.php/get-involved-pg [accessed 2014-05-13] [WebCite Cache ID 6PY2uelW9]
34. Gazmararian JA, Elon L, Yang B, Graham M, Parker R. Text4baby program: an opportunity to reach underserved pregnant and postpartum women? Matern Child Health J 2014 Jan;18(1):223-232. [doi: 10.1007/s10995-013-1258-1] [Medline: 23494485]
35. Prochaska JO, Velicer WF, Rossi JS, Goldstein MG, Marcus BH, Rakowski W, et al. Stages of change and decisional balance for 12 problem behaviors. Health Psychol 1994 Jan;13(1):39-46. [Medline: 8168470]
36. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot 1997;12(1):38-48. [Medline: 10170434]
37. Snyder LB, Hamilton MA, Mitchell EW, Kiwanuka-Tondo J, Fleming-Milici F, Proctor DA. A meta-analysis of the effect of mediated health communication campaigns on behavior change in the United States. J Health Commun 2004;9 Suppl 1:71-96. [doi: 10.1080/10810730490271548] [Medline: 14960405]
38. Evans WD. Social marketing campaigns and children's media use. Future Child 2008;18(1):181-203. [Medline: 21338010]

Abbreviations

BL: baseline
BRFSS: Behavioral Risk Factor Surveillance Survey
DOD: Department of Defense
FL: follow-up
GEE: generalized estimating equations
HBM: health belief model
KAB: prenatal knowledge, attitudes, and beliefs
LMP: last menstrual period
mHealth: mobile health
MHS: Military Health Service
MI: multiple imputation
NHANES: National Health and Nutritional Examination Survey
RCT: randomized controlled trial
SCT: social cognitive theory
TATRC: Telemedicine and Advanced Technology Research Center
TTM: transtheoretical model

Edited by G Eysenbach; submitted 03.02.14; peer-reviewed by J Gazmaranian, E Heatley; comments to author 14.03.14; revised version received 04.04.14; accepted 28.04.14; published 20.05.14.

Please cite as:
Evans WD, Wallace Bihm J, Szekely D, Nielsen P, Murray E, Abroms L, Snider J
Initial Outcomes From a 4-Week Follow-Up Study of the Text4baby Program in the Military Women's Population: Randomized Controlled Trial
J Med Internet Res 2014;16(5):e131
URL: http://www.jmir.org/2014/5/e131/
doi:10.2196/jmir.3297
PMID:24846909

Original Paper

Medical Students and Personal Smartphones in the Clinical Environment: The Impact on Confidentiality of Personal Health Information and Professionalism

Kim Tran[1], BSc; Dante Morra[2], MD, FRCPC, MBA; Vivian Lo[1], MASc; Sherman D Quan[3], MSc; Howard Abrams[1,4], MD, FRCPC; Robert C Wu[1,4], MD, FRCPC, MSc

[1]University Health Network, Centre for Innovation in Complex Care, Toronto, ON, Canada

[2]Trillium Health Partners, Mississauga, ON, Canada

[3]Trillium Health Partners, Institute for Better Health, Mississauga, ON, Canada

[4]University Health Network, Division of General Internal Medicine, Toronto, ON, Canada

Corresponding Author:
Kim Tran, BSc
University Health Network
Centre for Innovation in Complex Care
Toronto General Hospital, G-NU 403
Toronto, ON, M5G2C4
Canada
Phone: 1 416 340 4800 ext 5472
Fax: 1 416 595 5826
Email: kim.tran2@uhn.ca

Abstract

Background: Smartphones are becoming ubiquitous in health care settings. The increased adoption of mobile technology such as smartphones may be attributed to their use as a point-of-care information source and to perceived improvements in clinical communication and efficiency. However, little is known about medical students' use of personal smartphones for clinical work.

Objective: The intent of the study was to examine final-year medical students' experience with and attitudes toward using personal mobile technology in the clinical environment, with respect to the perceived impact on patient confidentiality and provider professionalism.

Methods: Cross-sectional surveys were completed by final-year medical students at the University of Toronto. Respondents were asked about the type of personal mobile phone they use, security features on their personal phone, experiences using their personal phone during clinical rotations, and attitudes about using their personal phone for clinical work purposes.

Results: The overall response rate was 45.4% (99/218). Smartphone ownership was prevalent (98%, 97/99) with the majority (86%, 85/99) of participants using their personal phones for patient-related communication during clinical rotations. A total of 26% (26/99) of participants reported not having any type of security feature on their personal phone, 94% (90/96) of participants agreed that using their personal phone for clinical work makes them more efficient, and 86% (82/95) agreed that their personal phone allows them to provide better patient care. Although 68% (65/95) of participants believe that the use of personal phones for patient-related communication with colleagues poses a risk to the privacy and confidentiality of patient health information, 22% (21/96) of participants still use their personal phone to text or email identifiable patient data to colleagues.

Conclusions: Our findings suggest that the use of personal smartphones for clinical work by medical students is prevalent. There is a need to more fully address the threat to patient confidentiality posed by the use of unsecured communication devices such as smartphones.

(J Med Internet Res 2014;16(5):e132) doi:10.2196/jmir.3138

KEYWORDS

medical informatics; communication; hospitals; mobile phone; smartphones

Introduction

Smartphones are becoming ubiquitous in the health care setting. The rise in mobile technology such as smartphones may be attributed to perceived improvements in clinical communication, efficiency, and clinical skills [1-8]. Effective communication between health care providers is vital for optimal patient care. The importance of effective communication in the delivery of care is evident given that communication failures represent the most common cause of preventable disability or death [9].

Smartphones are also being recognized for their use in medical education and training. With smartphones being described as a "learn anywhere" resource [10], medical students and doctors are using medical-related applications for both educational and clinical purposes. Popular applications include those for medication/drug reference, disease diagnosis/management, and clinical scoring systems [11].

Although some studies have shown multiple benefits associated with increased connectivity from smartphone use, negative consequences of its use have also been described. Distracted doctoring due to frequent smartphone interruptions can result in adverse events such as medical errors [12,13]. Increasing use of personal smartphones for clinical communication has been observed, possibly due to the lack of an existing secure and efficient hospital communication system [3,14]. In addition, the use of personal smartphones for communicating patient information and the potential for unprofessional behavior have been described [3,15]. Finally, there are privacy concerns for patient health information to be communicated through unsecure methods such as email and text using personal smartphones [3].

This study explores the uses of personal smartphones by medical students during their clinical rotations and describes the perceived impact on the confidentiality of personal health information and professionalism.

Methods

Study Population

Participants were fourth-year medical students from the Faculty of Medicine at the University of Toronto. Participants would have been exposed to all of their clinical rotations in the various medical specialties.

Survey Development

We developed the survey through an iterative process using standard survey methodology [16]. A literature search was conducted on MEDLINE to identify publications describing the uses of personal smartphones in the clinical environment (search terms: [cellular phone or smartphone or smart phone or iPhone or Android or BlackBerry or iPad or Windows mobile or personal digital assistant or mobile computer or mobile phone] AND [medical student or resident or physician] AND [medical education]). Semi-structured interviews were also conducted with seven medical students from the University of Toronto to examine their use of personal phones in the clinical environment. Important domains were identified and questions were generated through the literature review, interviews, and expert feedback. An expert group in the field of clinical informatics (RW, DM, VL, and SQ) reviewed the survey for content and face validity, comprehensiveness, and clarity. Pre-testing occurred with two focus groups consisting of individuals with research and/or design backgrounds who reviewed the survey for clarity and interpretation of individual questions. We then pilot-tested the survey with nine medical students and obtained feedback. The final survey consisted of 19 questions (Multimedia Appendix 1). A 5-point scale was used to express frequency for seven items and a 5-point Likert scale was used to express level of agreement for nine questions. The remaining three questions asked about the type, uses, and security features on medical students' personal mobile phones.

Data Collection and Analysis

In February 2013, medical students from the University of Toronto were surveyed during their final year of the medical school curriculum. Each student was provided with a paper survey at the beginning of his/her Transition to Residency course (all fourth-year medical students are required to take the course). A project manager for the Undergraduate Medical Education program distributed surveys at the beginning of class and completed surveys were collected during a class break. Students were informed that participation was voluntary and responses were anonymous and confidential. The study was approved by The University of Toronto Research Ethics Board. Descriptive statistics were generated from the survey results using Microsoft Excel.

Results

Uses of Personal Smartphones

The overall response rate was 45.4% (99/218). Nearly all (98%, 97/99) of the respondents currently owned a personal smartphone and the majority (79%, 78/99) of participants owned iPhones (Table 1).

Medical students reported using their personal smartphones for multiple purposes during their clinical rotations. The majority of students used their personal phone to communicate with medical team members about patient-related matters (86%, 85/99) and non−patient-related matters (93%, 92/99). Although 71% (70/99) of students had password protection on their phone, the survey revealed that 26% (26/99) of students' phones lacked any type of security feature.

Table 1. The type, uses, and security features on medical students' personal mobile phones (n=99).

Question	Answer options	n (%)
1. What type of personal mobile phone do you currently use?[a]	iPhone	78 (79)
	BlackBerry	6 (6)
	Windows Phone	0 (0)
	Android	14 (14)
	Cellular phone (non-smartphone)	2 (2)
	Other: Nokia smartphone	1 (1)
2. How do you use your personal mobile phone during clinical rotations?	Communication with patients	3 (3)
	Communication with other medical team members (patient-related)	85 (86)
	Communication with other medical team members (not patient-related)	92 (93)
	Medical references, resources, and applications	92 (93)
	View patient information	6 (6)
	Personal purposes (not work-related)	89 (90)
3. What type of security features do you have on your personal mobile phone?	Password protection	70 (71)
	Encryption	5 (5)
	I don't know	6 (6)
	None	26 (26)

[a]Two participants reported using two types of personal mobile phones.

The Disruptive Nature of Smartphones

A total of 46% (45/97) of medical students stated that they had answered/made a call, texted, or emailed on their personal phone during patient encounters (Table 2, Q1). However, 93% (89/96) of students perceived that their senior resident or attending physician interrupted patient meetings to answer/make a call, text, or email (Table 2, Q2). The disruptive nature of mobile phones also appeared to impact educational sessions with 31% (30/97) of medical students and 19% (18/96) of senior residents and attending physicians frequently interrupting an educational session to use their mobile phone (Table 2, Q3, Q4). In terms of personal use of their smartphones, 64% (61/95) of students frequently or always used their personal mobile phone for personal matters during their clinical rotations (Table 2, Q5).

Table 2. Participants' experiences using personal mobile technology during clinical rotations (n=99)[a].

Question	Never, n (%)	Rarely (1-3 times / month), n (%)	Occasionally (1-6 times / week), n (%)	Frequently (1-10 times / day), n (%)	Always (>10 times / day), n (%)
Q1. I have answered/made a call, texted, or emailed on my personal mobile phone while I was with a patient.	52 (54)	35 (36)	10 (10)	0 (0)	0 (0)
Q2. My senior resident or attending physician has interrupted a patient meeting to answer/make a call, text, or email.	7 (7)	38 (40)	41 (43)	10 (10)	0 (0)
Q3. I have answered/made a call, texted, or emailed on my personal mobile phone while I was in an educational session (eg, teaching rounds, bullet rounds, etc)	6 (6)	24 (25)	32 (33)	30 (31)	5 (5)
Q4. My senior resident or attending physician has interrupted an educational session to answer/make a call, text, or email.	3 (3)	41 (43)	34 (35)	18 (19)	0 (0)
Q5. I used my personal mobile phone for personal matters (eg, personal texts, calls, etc) during clinical rotations.	2 (2)	7 (7)	25 (26)	49 (52)	12 (13)
Q6. I used my personal mobile phone to text or email identifiable patient data (eg, patient last name, OHIP number, medical record number, etc) to colleagues.	75 (78)	17 (18)	3 (3)	1 (1)	0 (0)
Q7. My senior resident or attending physician has texted or emailed identifiable patient data to colleagues.	40 (44)	38 (42)	9 (10)	4 (4)	0 (0)

[a]A total of 99 surveys were returned but some participants did not answer every question.

Communicating Patients' Personal Health Information

In total, 78% (75/96) of students reported that they had never used their personal phone to text or email identifiable patient information to colleagues (Table 2, Q6). However, students reported that their senior residents or attending physicians were more likely to communicate identifiable patient information to colleagues, as only 44% (40/91) of students reported that their senior or attending had never texted or emailed identifiable patient information (Table 2, Q7). In terms of efficiency and patient care, 94% (90/96) of students believed that using their personal phone for clinical work made them more efficient and 86% (82/95) of students believed their personal phone allowed them to provide better patient care (Table 3, Q14, Q16). Although 68% (65/95) of students believed the use of personal phones for patient-related communication with colleagues poses a risk to the privacy and confidentiality of patient health information (Table 3, Q12), 22% (21/96) of participants still used their personal phone to text or email identifiable patient data to colleagues (Table 2, Q6). The majority of students (57%, 55/96) believed that the efficiency of communicating with colleagues through text and email using their personal phone outweighed the risk to the privacy and confidentiality of patient health information (Table 3, Q15).

Table 3. Participants' attitudes about using personal mobile technologies for clinical work purposes (n=99)[a].

Question	Strongly disagree, n (%)	Disagree, n (%)	Neutral, n (%)	Agree, n (%)	Strongly agree, n (%)
Q8. The medical school curriculum has educated me on appropriate and inappropriate ways to use my personal mobile phone for communicating patient information.	3 (3)	18 (19)	18 (19)	48 (51)	8 (8)
Q9. My senior resident or attending physician has given me feedback on appropriate and inappropriate ways to use my personal mobile phone for communicating patient information.	22 (23)	36 (37)	16 (17)	19 (20)	3 (3)
Q10. The medical school curriculum has educated me on appropriate and inappropriate ways to conduct myself professionally with mobile technology.	5 (5)	23 (24)	27 (28)	36 (38)	5 (5)
Q11. My senior resident or attending physician has given me feedback on appropriate and inappropriate ways to conduct myself professionally with mobile technology.	19 (20)	42 (44)	26 (27)	9 (9)	0 (0)
Q12. The use of personal mobile phones for patient-related communication with colleagues poses a risk to the privacy and confidentiality of patient health information.	2 (2)	7 (7)	21 (22)	38 (40)	27 (28)
Q13. My personal mobile phone is distracting during clinical work.	17 (18)	40 (42)	20 (21)	19 (20)	0 (0)
Q14. Using my personal mobile phone for clinical work makes me more efficient.	0 (0)	1 (1)	5 (5)	54 (56)	36 (38)
Q15. The efficiency of communicating with colleagues through text and email using my personal mobile phone outweighs the risk to the privacy and confidentiality of patient health information.	5 (5)	12 (13)	24 (25)	46 (48)	9 (9)
Q16. Using my personal mobile phone for clinical work allows me to provide better patient care.	0 (0)	0 (0)	13 (14)	59 (62)	23 (24)

[a]A total of 99 surveys were returned but some participants did not answer every question.

Preparedness for Using Personal Smartphones in a Clinical Environment

A total of 59% (56/95) of students agreed or strongly agreed that their medical school curriculum had educated them on appropriate and inappropriate ways to use their personal mobile phone for communicating patient information (Table 3, Q8); 43% (41/96) of students believed their medical school curriculum had educated them on appropriate and inappropriate ways to conduct themselves professionally with mobile technology (Table 3, Q10).

Discussion

Principal Results

Personal smartphone use among medical students has become ubiquitous in health care settings. Our results provide a description of how and why medical students are using their personal smartphones. In addition, we describe the possible issues that could arise relating to medical students' level of preparedness on the appropriate and inappropriate use of their smartphones in the clinical environment. Students are using their personal smartphones for work-related functions such as communicating with medical team members about patient-related and non−patient-related matters and using medical references, resources, and applications. They perceive that smartphone use increases their efficiency. While they communicate patient-related information using their personal phones, most medical students did not report communicating patient identifiable personal health information (PHI) in texts or emails. However, the majority (56%) of students reported that their senior residents and attending physicians had communicated patient identifiable PHI. In terms of preparedness, approximately half of students perceived they were educated on appropriate uses of their personal smartphones.

The personal smartphones of most participants lacked the necessary security features to protect the sensitive information that they may be sharing. As required by the Personal Health Information Protection Act (PHIPA), smartphones must be configured to operate in a secure manner when used to transmit or store personal health information [17]. Security features include the encryption of transmissions, password protection, and automated data wiping [17,18]. In recent years, the US Department of Health and Human Services has issued large fines to health care organizations and groups violating policies set out in the Health Insurance Portability and Accountability Act (HIPAA) [19]. These actions present a clear message that all health care providers and organizations will be held accountable for protecting their patients' health information.

Despite security concerns over using personal smartphones for clinical work purposes, medical students perceive that their devices make them more efficient and allow them to provide better patient care. The majority believe that the benefits of

perceived better care outweigh the possible harms of unsecure communication. However, this increased connectivity may have a negative impact on professionalism such as "distracted doctoring", which may disrupt patient care and education.

Although a vast majority of medical students are using their personal smartphones in the clinical environment, many students do not feel that the medical school curriculum or role modeling has educated them on appropriate and inappropriate ways to use their personal smartphone for clinical work. There is increasing recognition that smartphone use by clinicians can be perceived to be unprofessional [2]. By answering their phone or responding to a text message during patient encounters, medical students and physicians can be perceived to be rude [20]. Through the medical school curriculum and role modeling, mobile etiquette should be taught to students so that they know where, when, and how it is appropriate to use their mobile technologies. Institutional policies regarding smartphone use in the clinical environment may also be beneficial [21]. This education would address issues of professionalism that can arise with the use of personal smartphones in clinical environments [12,13,15].

Our findings raise concerns over the security of personal health information. The use of personal smartphones for clinical work may increase efficiency, but there is concern about privacy breaches through unsecure sharing of confidential information. While individual clinicians including medical students, residents, and staff physicians need to understand the importance of keeping personal health information secure, it is the responsibility of the institutions to provide an effective, secure communication infrastructure for clinicians. Otherwise, we can expect ongoing privacy breaches.

Limitations

This study has several limitations. The study only included medical students from a single university in Canada and our response rate was only 45.4%. We may also have a biased selection of medical students who own smartphones. However, the university is affiliated with five academic teaching hospitals. With a total of 99 responses, we believe that these results are likely to be generalizable to medical teaching institutions in Canada and the United States. The study also examined self-reported experiences. Participant responses were, therefore, subject to both recall bias and response bias. In addition, due to the sensitive nature of PHI, communication of PHI may be under-reported. Our findings show that students perceived their senior residents or attending physicians to communicate PHI more often than they reported for themselves. However, it is unclear whether the devices their senior residents or attending physicians were using to communicate PHI were personal phones or institutional devices.

Comparison With Prior Work

Some of our findings are consistent with the literature reporting high use of smartphones by clinicians for work purposes. Consistent with previous literature, smartphones are being used for clinical and educational purposes and perceived improvements in efficiency have been reported [1-4,6]. Issues around interruptions resulting from smartphone use have also previously been described [22-24]. Additionally, professionalism issues have been described with medical trainees using smartphones in the clinical setting [20]. However, our study contributes original knowledge regarding *personal* smartphone use by medical students in the clinical environment: specifically, the prevalence of smartphone ownership among medical students, the various uses of personal smartphones, and students' level of preparedness for using their personal smartphones in the clinical environment.

Conclusions

The use of personal smartphones in the clinical environment is an established reality. It is evident that medical students prefer to use their personal smartphones for clinical work as they perceive that these devices make them more efficient and allow them to provide better patient care. With the popularity of personal smartphones, it is critical that more attention be focused on educating medical professionals on how to appropriately use their personal devices for clinical work as well as adopting secure means for clinical communication.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Smartphone survey.

[PDF File (Adobe PDF File), 529KB - jmir_v16i5e132_app1.pdf]

References

1. Wu RC, Morra D, Quan S, Lai S, Zanjani S, Abrams H, et al. The use of smartphones for clinical communication on internal medicine wards. J Hosp Med 2010 Dec;5(9):553-559. [doi: 10.1002/jhm.775] [Medline: 20690190]
2. Wu R, Rossos P, Quan S, Reeves S, Lo V, Wong B, et al. An evaluation of the use of smartphones to communicate between clinicians: a mixed-methods study. J Med Internet Res 2011;13(3):e59 [FREE Full text] [doi: 10.2196/jmir.1655] [Medline: 21875849]

3. Wu RC, Lo V, Morra D, Wong BM, Sargeant R, Locke K, et al. The intended and unintended consequences of communication systems on general internal medicine inpatient care delivery: a prospective observational case study of five teaching hospitals. J Am Med Inform Assoc 2013;20(4):766-777. [doi: 10.1136/amiajnl-2012-001160] [Medline: 23355461]
4. Patel BK, Chapman CG, Luo N, Woodruff JN, Arora VM. Impact of mobile tablet computers on internal medicine resident efficiency. Arch Intern Med 2012 Mar 12;172(5):436-438. [doi: 10.1001/archinternmed.2012.45] [Medline: 22412110]
5. Davis JS, Garcia GD, Wyckoff MM, Alsafran S, Graygo JM, Withum KF, et al. Use of mobile learning module improves skills in chest tube insertion. J Surg Res 2012 Sep;177(1):21-26. [doi: 10.1016/j.jss.2012.03.022] [Medline: 22487392]
6. Wallace S, Clark M, White J. 'It's on my iPhone': attitudes to the use of mobile computing devices in medical education, a mixed-methods study. BMJ Open 2012;2(4):1 [FREE Full text] [doi: 10.1136/bmjopen-2012-001099] [Medline: 22923627]
7. Tews M, Brennan K, Begaz T, Treat R. Medical student case presentation performance and perception when using mobile learning technology in the emergency department. Med Educ Online 2011;16:1 [FREE Full text] [doi: 10.3402/meo.v16i0.7327] [Medline: 22013378]
8. Clay CA. Exploring the use of mobile technologies for the acquisition of clinical skills. Nurse Educ Today 2011 Aug;31(6):582-586. [doi: 10.1016/j.nedt.2010.10.011] [Medline: 21112132]
9. Wilson RM, Runciman WB, Gibberd RW, Harrison BT, Newby L, Hamilton JD. The Quality in Australian Health Care Study. Med J Aust 1995 Nov 6;163(9):458-471. [Medline: 7476634]
10. Trelease RB. Diffusion of innovations: smartphones and wireless anatomy learning resources. Anat Sci Educ 2008;1(6):233-239. [doi: 10.1002/ase.58] [Medline: 19109851]
11. Payne KB, Wharrad H, Watts K. Smartphone and medical related App use among medical students and junior doctors in the United Kingdom (UK): a regional survey. BMC Med Inform Decis Mak 2012;12:121 [FREE Full text] [doi: 10.1186/1472-6947-12-121] [Medline: 23110712]
12. Ross S, Forgie S. Distracted doctoring: smartphones before patients? CMAJ 2012 Sep 4;184(12):1440 [FREE Full text] [doi: 10.1503/cmaj.120462] [Medline: 22891206]
13. Westbrook JI, Woods A, Rob MI, Dunsmuir WT, Day RO. Association of interruptions with an increased risk and severity of medication administration errors. Arch Intern Med 2010 Apr 26;170(8):683-690. [doi: 10.1001/archinternmed.2010.65] [Medline: 20421552]
14. Wu RC, Tran K, Lo V, O'Leary KJ, Morra D, Quan SD, et al. Effects of clinical communication interventions in hospitals: a systematic review of information and communication technology adoptions for improved communication between clinicians. Int J Med Inform 2012 Nov;81(11):723-732. [doi: 10.1016/j.ijmedinf.2012.05.014] [Medline: 22727613]
15. Robinson T, Cronin T, Ibrahim H, Jinks M, Molitor T, Newman J, et al. Smartphone use and acceptability among clinical medical students: a questionnaire-based study. J Med Syst 2013 Jun;37(3):9936. [doi: 10.1007/s10916-013-9936-5] [Medline: 23526127]
16. Burns KE, Duffett M, Kho ME, Meade MO, Adhikari NK, Sinuff T, ACCADEMY Group. A guide for the design and conduct of self-administered surveys of clinicians. CMAJ 2008 Jul 29;179(3):245-252 [FREE Full text] [doi: 10.1503/cmaj.080372] [Medline: 18663204]
17. Cavoukian A, Rossos PG. Information and Privacy Commissioner of Ontario. Personal health information: a practical tool for physicians transitioning from paper-based records to electronic health records URL: http://www.ipc.on.ca/images/Resources/phipa-toolforphysicians.pdf [accessed 2013-11-21] [WebCite Cache ID 6LIuKqbBr]
18. Farnan JM, Snyder Sulmasy L, Worster BK, Chaudhry HJ, Rhyne JA, Arora VM, American College of Physicians Ethics, Professionalism and Human Rights Committee, American College of Physicians Council of Associates, Federation of State Medical Boards Special Committee on Ethics and Professionalism*. Online medical professionalism: patient and public relationships: policy statement from the American College of Physicians and the Federation of State Medical Boards. Ann Intern Med 2013 Apr 16;158(8):620-627. [doi: 10.7326/0003-4819-158-8-201304160-00100] [Medline: 23579867]
19. U.S. Department of Health & Human Services. Annual Report to Congress on HIPAA Privacy Rule and Security Rule Compliance For Calendar Years 2009 and 2010 URL: http://www.hhs.gov/ocr/privacy/hipaa/enforcement/complianceрept.pdf [accessed 2013-11-21] [WebCite Cache ID 6LIuqmMCE]
20. Wu RC, Tzanetos K, Morra D, Quan S, Lo V, Wong BM. Educational impact of using smartphones for clinical communication on general medicine: more global, less local. J Hosp Med 2013 Jul;8(7):365-372. [doi: 10.1002/jhm.2037] [Medline: 23713054]
21. Katz-Sidlow RJ, Lindenbaum Y, Sidlow R. Residents' attitudes toward a smartphone policy for inpatient attending rounds. J Hosp Med 2013 Sep;8(9):541-542. [doi: 10.1002/jhm.2070] [Medline: 23955894]
22. Katz-Sidlow RJ, Ludwig A, Miller S, Sidlow R. Smartphone use during inpatient attending rounds: prevalence, patterns and potential for distraction. J Hosp Med 2012 Oct;7(8):595-599. [doi: 10.1002/jhm.1950] [Medline: 22744793]
23. Wu RC, Tzanetos K, Morra D, Quan S, Lo V, Wong BM. Educational impact of using smartphones for clinical communication on general medicine: more global, less local. J Hosp Med 2013 Jul;8(7):365-372. [doi: 10.1002/jhm.2037] [Medline: 23713054]
24. Edwards A, Fitzpatrick LA, Augustine S, Trzebucki A, Cheng SL, Presseau C, et al. Synchronous communication facilitates interruptive workflow for attending physicians and nurses in clinical settings. Int J Med Inform 2009 Sep;78(9):629-637. [doi: 10.1016/j.ijmedinf.2009.04.006] [Medline: 19482544]

Abbreviations

PHI: personal health information

Edited by G Eysenbach; submitted 27.11.13; peer-reviewed by R Katz, R Sidlow, A Peña-Oscuvilca; comments to author 11.02.14; revised version received 28.03.14; accepted 27.04.14; published 22.05.14.

Please cite as:
Tran K, Morra D, Lo V, Quan SD, Abrams H, Wu RC
Medical Students and Personal Smartphones in the Clinical Environment: The Impact on Confidentiality of Personal Health Information and Professionalism
J Med Internet Res 2014;16(5):e132
URL: http://www.jmir.org/2014/5/e132/
doi:10.2196/jmir.3138
PMID:24855046

Original Paper

Daily Collection of Self-Reporting Sleep Disturbance Data via a Smartphone App in Breast Cancer Patients Receiving Chemotherapy: A Feasibility Study

Yul Ha Min[1], RN, PhD; Jong Won Lee[2,3], MD; Yong-Wook Shin[4], MD, PhD; Min-Woo Jo[5], MD, PhD; Guiyun Sohn[2], MD; Jae-Ho Lee[6], MD, PhD; Guna Lee[7], RN, MPH; Kyung Hae Jung[8], MD, PhD; Joohon Sung[3], MD, PhD; Beom Seok Ko[2], MD, PhD; Jong-Han Yu[2], MD; Hee Jeong Kim[2], MD, PhD; Byung Ho Son[2], MD, PhD; Sei Hyun Ahn[2], MD, PhD

[1]Seoul National University College of Nursing, Seoul, Republic of Korea

[2]Department of Surgery, University of Ulsan College of Medicine, Asan Medical Center, Seoul, Republic of Korea

[3]Department of Epidemiology, School of Public Health and Institute of Health and Environment, Seoul National University, Seoul, Republic of Korea

[4]Department of Psychiatry, University of Ulsan College of Medicine, Asan Medical Center, Seoul, Republic of Korea

[5]Department of Preventive Medicine, University of Ulsan College of Medicine, Asan Medical Center, Seoul, Republic of Korea

[6]Department of Emergency Medicine, University of Ulsan College of Medicine, Asan Medical Center, Seoul, Republic of Korea

[7]Medical Information Administration Team, Asan Medical Center, Seoul, Republic of Korea

[8]Department of Oncology, University of Ulsan College of Medicine, Asan Medical Center, Seoul, Republic of Korea

Corresponding Author:
Jong Won Lee, MD
Department of Surgery
University of Ulsan College of Medicine
Asan Medical Center
88 Olympic-ro 43-gil, Songpa-gu
Seoul, 138-736
Republic of Korea
Phone: 82 230105603
Fax: 82 24749027
Email: jongwonlee116@gmail.com

Abstract

Background: Improvements in mobile telecommunication technologies have enabled clinicians to collect patient-reported outcome (PRO) data more frequently, but there is as yet limited evidence regarding the frequency with which PRO data can be collected via smartphone applications (apps) in breast cancer patients receiving chemotherapy.

Objective: The primary objective of this study was to determine the feasibility of an app for sleep disturbance-related data collection from breast cancer patients receiving chemotherapy. A secondary objective was to identify the variables associated with better compliance in order to identify the optimal subgroups to include in future studies of smartphone-based interventions.

Methods: Between March 2013 and July 2013, patients who planned to receive neoadjuvant chemotherapy for breast cancer at Asan Medical Center who had access to a smartphone app were enrolled just before the start of their chemotherapy and asked to self-report their sleep patterns, anxiety severity, and mood status via a smartphone app on a daily basis during the 90-day study period. Push notifications were sent to participants daily at 9 am and 7 pm. Data regarding the patients' demographics, interval from enrollment to first self-report, baseline Beck's Depression Inventory (BDI) score, and health-related quality of life score (as assessed using the EuroQol Five Dimensional [EQ5D-3L] questionnaire) were collected to ascertain the factors associated with compliance with the self-reporting process.

Results: A total of 30 participants (mean age 45 years, SD 6; range 35-65 years) were analyzed in this study. In total, 2700 daily push notifications were sent to these 30 participants over the 90-day study period via their smartphones, resulting in the collection of 1215 self-reporting sleep-disturbance data items (overall compliance rate=45.0%, 1215/2700). The median value of individual patient-level reporting rates was 41.1% (range 6.7-95.6%). The longitudinal day-level compliance curve fell to 50.0% at day 34 and reached a nadir of 13.3% at day 90. The cumulative longitudinal compliance curve exhibited a steady decrease

by about 50% at day 70 and continued to fall to 45% on day 90. Women without any form of employment exhibited the higher compliance rate. There was no association between any of the other patient characteristics (ie, demographics, and BDI and EQ5D-3L scores) and compliance. The mean individual patient-level reporting rate was higher for the subgroup with a 1-day lag time, defined as starting to self-report on the day immediately after enrollment, than for those with a lag of 2 or more days (51.6%, SD 24.0 and 29.6%, SD 25.3, respectively; *P*=.03).

Conclusions: The 90-day longitudinal collection of daily self-reporting sleep-disturbance data via a smartphone app was found to be feasible. Further research should focus on how to sustain compliance with this self-reporting for a longer time and select subpopulations with higher rates of compliance for mobile health care.

(J Med Internet Res 2014;16(5):e135) doi:10.2196/jmir.3421

KEYWORDS

mobile applications; self report; compliance; breast cancer

Introduction

Electronic health (eHealth) can be defined as the practice of medicine and public health using information and communication technology (ICT), such as computers, mobile phones, and satellite communications [1]. The term mobile health (mHealth) has emerged as a subcategory of eHealth and is now used when the practice involves using wireless communications and especially mobile phones or smartphones, which are now used by more than 70% of the population in some countries, including the United Arab Emirates, South Korea, Saudi Arabia, and Singapore (as of 2013; see Multimedia Appendix 1). Although evidence regarding the value of incorporating mobile phone-based patient-reported outcome (PRO) data into cancer patients' care is still in the embryonic stage, noteworthy to date are demonstrations of the feasibility of mHealth—and specifically the use of mobile phones—to assist in the collection of PROs on disease-related vital signs [2,3], treatment-related side effects [4], and possibly comprehensive psychological status [5].

Prior to the eHealth era, the collection of PROs in the oncology field was based simply on patient recall at clinic visits or by asking patients to keep paper diaries outside the clinic. However, patient recall is inherently inaccurate and plagued by potential bias and the actual compliance with keeping paper diaries according to protocols has proved to be much lower than was expected [6], thus undermining the rationale for using it. The advent of ICT, which can build survey systems with a broad range of clinical uses, has made it possible to capture PRO data in real time and broken through a key barrier to the use PRO data in the clinical-care setting. These kinds of electronic PROs (ePROs) from waiting rooms in the hospital [7,8], as well as from home [9], have reportedly been successfully collected with high mean compliance rates and patient satisfaction and are now increasingly used in routine outpatient cancer care to guide clinical decisions and enhance communication. However, most of the available evidence has been obtained from studies using tablets or desktop computers [7-9] and none of the evidence collected using mobile phones [2,4,10-12] involved the currently available smartphones that have unique functions such as push notification and flexible applications (apps). To the best of our knowledge, there is limited information regarding the feasibility and acceptability of smartphone-based collection of PROs for cancer patients receiving chemotherapy and in particular for daily data collections over long periods of time.

The primary objective of this study was to determine the feasibility of using a smartphone app to collect sleep disturbance-related data from breast cancer patients receiving chemotherapy, with overall and individual reporting rates. We also sought to determine whether the patients stopped responding via the app after a short time by calculating a longitudinal reporting rate per day over a period of 90 days. A secondary objective was to elucidate the variables associated with a higher compliance rate in order to identify the optimal subgroups to include in future studies of smartphone-based interventions.

Methods

Participant Recruitment

Patients who were planning to receive neoadjuvant chemotherapy for breast cancer were recruited for participation in this feasibility study when they were admitted to the Breast Cancer Center, Asan Medical Center, Seoul, South Korea, for 2 nights to evaluate the disease status for preoperative chemotherapy between March 2013 and July 2013. To be eligible for this study, patients had to indicate app avidity, be current iOS or Android smartphone users, and be able to read and understand Korean. In some cases, the app did not run properly due to unexpected incompatibility between the smartphones' display specifications and the app.

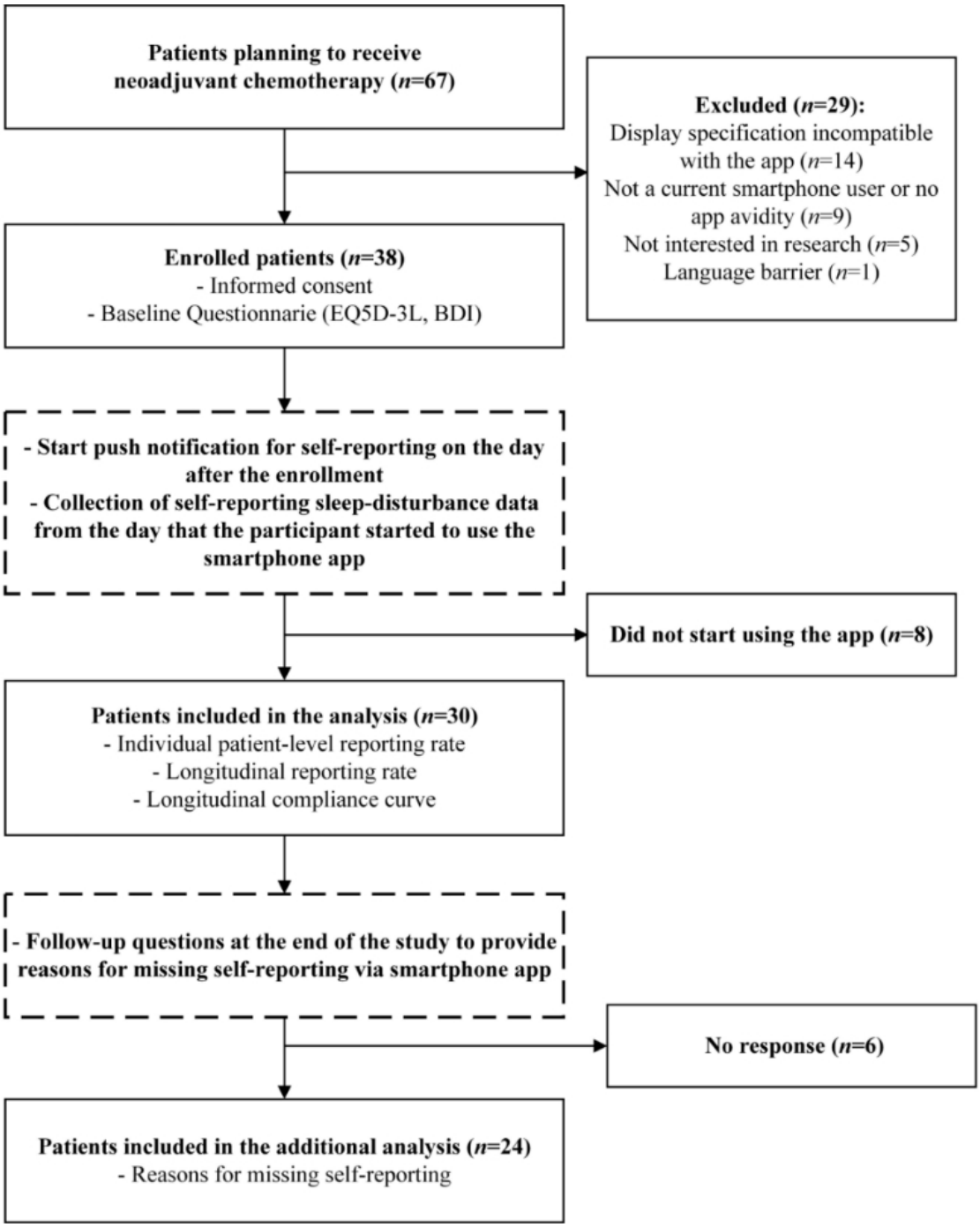
Among the 67 patients initially recruited for participation, 29 were excluded for the following reasons: incompatibility of the app with their smartphone's display specifications (n=14), not a current smartphone user or no app avidity (n=9), not interested in the research (n=5), and language barrier (n=1). A total of 38 patients were enrolled and provided informed consent to participate (IRB no. 2012-0709). Of these, 8 patients were not included in the analysis performed in this feasibility study because they did not start to use the app or did not report sleep-disturbance symptoms at all after providing consent to participate at admission (Figure 1).

The participants were presented with information regarding the study and the app, called "Pit-a-Pat", which was developed for both the iOS and Android platforms, and has been described previously [13]. A 30-minute interview with each patient at

admission included the following items: downloading of the app; entering baseline demographics such as age, educational attainment, marital status, cohabitation status, and occupation into the app; instructions on how to log in to the app and report information about sleep disturbance (possibly caused by the breast cancer itself and the chemotherapy); and baseline survey of health-related quality of life (HRQOL) status as assessed using the EQ5D-3L questionnaire and depression status as assessed using Beck's Depression Inventory (BDI).

After a 2-night admission for workup, the participants were discharged and scheduled to visit the outpatient clinic 7 days later (on average) to receive neoadjuvant chemotherapy. The workup results for neoadjuvant chemotherapy, such as regional lymph node metastasis, hormone receptor expression, and planned chemotherapeutic agents regimens, were also collected.

Figure 1. Study design and participant flow.



Daily Collection of Symptoms via the App

Briefly, the app in this study was developed for cancer patients to self-report three kinds of health experience that may be caused by the effect of the diagnosis with breast cancer itself and the subsequent treatments: (1) sleep-disturbance symptoms related to mild depression, (2) acute symptoms related to cytotoxic chemotherapeutic agents, and (3) medication diary for antihormonal treatment such as tamoxifen and aromatase inhibitor. The sleep-disturbance symptoms related to mild

depression were assessed using a six-item questionnaire: (1) time of falling asleep, (2) time of waking up, (3) number of awakenings during the total sleep period, (4) quality of total sleep on a 10-point rating scale, (5) present mood status on a 7-point rating scale, and (6) severity of present anxiety on a 10-point rating scale. The sleeping and waking times (in minutes) were stored as continuous variables. The 10-point rating scale for quality of total sleep and severity of present anxiety was displayed as a distress thermometer and the 7-point rating scale for mood status was displayed as faces with a range of expressions. The six items were displayed in three sequential pages as shown in Figure 2.

Participants were informed that they could start reporting their symptoms via the app at any time from the day following their provision of consent. Once the app was activated, push notifications were sent to the participants daily at 9 am and 6 pm, asking them to daily self-report sleep patterns, anxiety severity, and mood status via the smartphone app over a 90-day period. After entering appropriate answers to all six items without missing values, a participant could touch the send button and be considered as having completed a self-report of her sleep-disturbance symptoms on that specific day. The same push notification process was repeated for all participants during the study period, irrespective of their reporting rates. A database system collected the information about daily requests and inputs (ie, anonymized user ID, item ID, date and time of input, and input value), as well as baseline demographics.

The database system was connected to our electronic medical record system to enable clinicians to review the data. During the 90-day study period, clinicians could review the participants' self-reporting data at the outpatient clinic at 3-week intervals, and managed them in the same way as the nonparticipating patients for the reported symptoms, since this was a study only related to feasibility. Participants were not provided with a smartphone for this study, instead being required to use their own device. Furthermore, they did not receive any reimbursement of expenses for extra data usage caused by utilization of the app, or any financial incentive for participating in the study.

At the end of the study period, the participants received a push notification requesting reasons for missing self-reporting, by choosing from the following possible responses: "The app didn't work properly", "I forgot", "I didn't think it was useful", "I was too sick", "I didn't feel like it", "It was inconvenient", or "I was too busy". Among the 30 participants who received 90 requests for self-reporting, 24 responded and their data were analyzed.

Figure 2. Screenshot of the app for self-reporting of sleep disturbance data.

Statistical Analysis

The feasibility of this self-reporting system was analyzed by calculating individual patient-level reporting rates, defined as the total number of days in which self-reporting was completed divided by 90 for each participant, and a longitudinal day-level reporting rate, defined as the total number of participants who completed self-reporting divided by 30 on a single specific day

from the start of self-reporting. A cumulative longitudinal day-level reporting rate was also defined as the mean of all longitudinal day-level reporting rates from the starting day until the specified day. For example, the longitudinal day-level reporting rate at day 34 indicates the number of participants self-reporting on the 34th day since the starting day divided by 30 and the cumulative longitudinal day-level reporting rate at day 34 indicates the mean value of all longitudinal day-level reporting rates from day 1 to day 34. A longitudinal compliance curve was plotted by connecting the 90 longitudinal day-level reporting rates for each day over the 90-day study period and a cumulative longitudinal compliance curve with 90 cumulative longitudinal day-level reporting rates over time. For the cumulative dropout probability curve, a dropout event was defined as when a cumulative longitudinal day-level reporting rate of a participant decreases first to below 28.6%, which means that she reported less than twice per week, which is almost the same as the weekly biodata collection.

For a priori threshold criteria for good compliance, prior feasibility studies in related contexts were reviewed [3,4,7-9,14] to determine whether this daily data-collection approach via smartphone app merited further development. On a reasonable threshold for feasibility, experts came to the conclusion that the previous evidence was insufficient because of small samples [3,4], use of a tablet, desktop computer, or conventional mobile phones rather than a current smartphone [3,4,7-9,14], and inclusion of patients with noncancerous chronic diseases [3]. Instead, we decided to explore the feasibility ourselves, focusing on the following characteristics of this study cohort: women with breast cancer, receiving chemotherapy, daily collection, smartphone with an app, and South Korea's high degree of smartphone penetration. In addition, we planned to establish an appropriate cutoff for the higher compliance rate after describing the distribution of various reporting rates.

The chi-square test and *t* test were used to identify differences in variables between groups with higher and lower rates of compliance. Baseline patient variables of interest included age, educational attainment, occupation status, HRQOL status, BDI score, and interval from enrollment to first self-reporting. This study was performed in a cohort with uniform characteristics in terms of race, cancer type and stage, and Eastern Cooperative Oncology Group performance status. Except where stated otherwise, the data are presented as mean (SD) values and the threshold of statistical significance was set at $P<.05$. All statistical analyses were performed using SPSS version 12.0 (SPSS, Chicago, IL, USA).

Results

Participant Characteristics

Between March 2013 and July 2013, 67 patients were initially eligible for participation. However, 14 patients (20%, 14/67) could not participate in this study because of inherent technological limitations in the development of the app causing incompatibility with their smartphone's display specifications. Among the 53 women who did not experience technical problems in using the app, 9 (17%, 9/53) showed no interest in either their smartphone or the app and 5 (9%, 5/53) refused to participate in this study because they were not interested in the research. Of the 38 eligible patients who were actually enrolled into this feasibility study, 8 (21%, 8/38) did not start to use the app. The final analysis thus included 30 patients (mean age 45 years, SD 6; range 36-65 years), all 30 of whom were Korean women who had been diagnosed with breast cancer within 4 weeks prior to enrollment; 73% (22/30) were 50 years old, 47% (14/30) had an educational attainment of college level or higher, and 43% (13/30) were currently employed. The baseline BDI and EQ5D-3L scores were 11.5 (SD 8.8) (range 0-35) and 0.92 (SD 0.09) (range 0.56-1.00), respectively. For tumor characteristics, lymph-node metastases were confirmed in 77% (23/30), 60% (18/30) were hormone-responsive breast cancer, and those with any kind of distant metastases were excluded. During the 90-day study period, all participants received adriamycin-based (87%, 26/30) or epirubicin-based (13%, 4/30) combinational chemotherapeutic agents, of whom 77% (23/30) received additional docetaxel because of positive lymph-node metastasis after this feasibility study (Table 1).

Table 1. Baseline demographics of patients (n=30).

Characteristic		Total (n=30) n (%)	Lower compliance rate (n=15)	Higher compliance rate (n=15)	*P*
Age					
Mean (SD) years		45 (6)	46 (5)	45 (8)	NS[a]
Range, years		36-65	38-55	36-65	
≤49 years		22 (73)	11 (73)	11 (73)	NS
≥50 years		8 (27)	4 (27)	4 (27)	
Level of educational attainment					NS
Up to high school		16 (53)	10 (67)	6 (40)	
College or greater		14 (47)	5 (33)	9 (60)	
Marital status					NS
Married		26 (87)	13 (87)	13 (86)	
Single		1 (3)	0 (0)	1 (7)	
Divorced		3 (10)	2 (13)	1 (7)	
Cohabiting					NS
Yes		28 (93)	15 (100)	13 (87)	
No		2 (7)	0 (0)	2 (13)	
Occupation					.03
Yes (of any kind)		13 (43)	10 (67)	3 (20)	
No (eg, stay-at-home mother)		17 (57)	5 (33)	12 (80)	
BDI [b] score					
Mean (SD)		11.5 (8.8)	11.9 (8.5)	11.1 (9.4)	NS
Range		0-35	1-33	0-35	
≤15		21 (70)	10 (67)	11 (73)	NS
≥16		9 (30)	5 (33)	4 (27)	
HRQOL [c] with EQ5D-3L [d]					
EQ5D-3L VAS [e] score					NS
	Mean (SD)	69.6 (16.1)	68.3 (16.0)	70.9 (16.6)	
	Range	38-99	45-95	38-99	
EQ5D-3L utility score					NS
	Mean (SD)	0.92 (0.09)	0.92 (0.07)	0.91 (0.11)	
	Range	0.56-1.00	0.74-1.00	0.56-1.00	
Mobility					NS
	No problems	29 (97)	15 (100)	14 (93)	
	Problems	1 (3)	0 (0)	1 (7)	
Self-care					NS
	No problems	30 (100)	15 (100)	15 (100)	
	Problems	0 (0)	0 (0)	0 (0)	
Usual activities					NS
	No problems	29 (97)	15 (100)	14 (93)	
	Problems	1 (3)	0 (0)	1 (7)	

Characteristic		Total (n=30) n (%)	Lower compliance rate (n=15)	Higher compliance rate (n=15)	*P*
Pain/discomfort					NS
	No problems	19 (63)	11 (73)	8 (53)	
	Problems	11 (37)	4 (27)	7 (47)	
Anxiety/depression					NS
	No problems	14 (47)	6 (40)	8 (53)	
	Problems	16 (53)	9 (60)	7 (47)	
Lymph node metastasis					NS
Negative		7 (23)	4 (27)	3 (20)	
Positive		23 (77)	11 (73)	12 (80)	
Hormone receptor status					NS
Positive		18 (60)	6 (40)	12 (80)	
Negative		12 (40)	9 (60)	3 (20)	
Neoadjuvant chemotherapy regimen					NS
AC[f]#4		7 (23)	4 (27)	3 (20)	
AC#4 followed by docetaxel#4		19 (64)	7 (46)	12 (80)	
FEC[g]#3 followed by docetaxel#3		4 (13)	4 (27)	0 (0)	

[a]Not significant.

[b]BDI: Beck's Depression Inventory.

[c]HRQOL: health-related quality of life.

[d]EQ5D-3L: EuroQol Five Dimensional Questionnaire.

[e]Visual analog scale.

[f]Combination of doxorubicin+cyclophosphamide chemotherapy.

[g]Combination of 5-fluorouracil+epirubicin+cyclophosphamide chemotherapy.

Feasibility Analysis

In total, 2700 daily push notifications were sent to the 30 participants via their smartphones during the 30-day study period; 1215 responses were received containing completed self-reporting sleep-disturbance data (overall compliance rate=45.0%; 1215/2700). The median individual patient-level reporting rates from the 30 participants was 41.1% (range 6.7-95.6%), with the following distribution: 6 (20%, 6/30), 9 (30%, 9/30), 6 (20%, 6/30), 7 (23%, 7/30), and 2 (7%, 2/30) ranging from 0% to 20%, 21% to 40%, 41% to 60%, 61% to 80%, and 81% to 100%, respectively (Figure 3).

The longitudinal day-level reporting rate on day 1 is theoretically 100%, because that day is the first day when all participants started to use the app. The longitudinal day-level reporting rates from day 1 to day 90 were calculated at daily intervals and plotted as a longitudinal compliance curve. The longitudinal compliance curve for this app decreased rapidly to about 50.0% at day 34 after the start of self-reporting, and continued to decrease steadily thereafter, reaching a nadir of 13.3% at day 90. The cumulative longitudinal compliance curve (representing overall compliance) revealed a steady decrease by about 50% at day 70, and continued to fall to 45% on day 90. Cumulative dropout probabilities showed that 10% (3/30) of all participants self-reported less than twice per week on average (cumulative longitudinal compliance rate <28.6%) during the first 28 days, 20% (6/30) during the first 56 days, and 33.3% (10/30) during the 90-day study period (Figure 4).

Figure 3. Distribution of individual patient-level reporting rates.

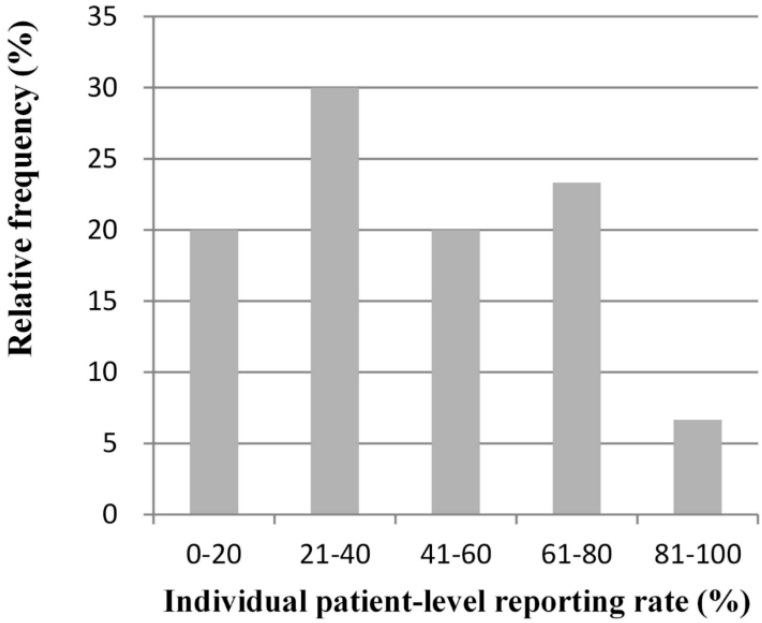
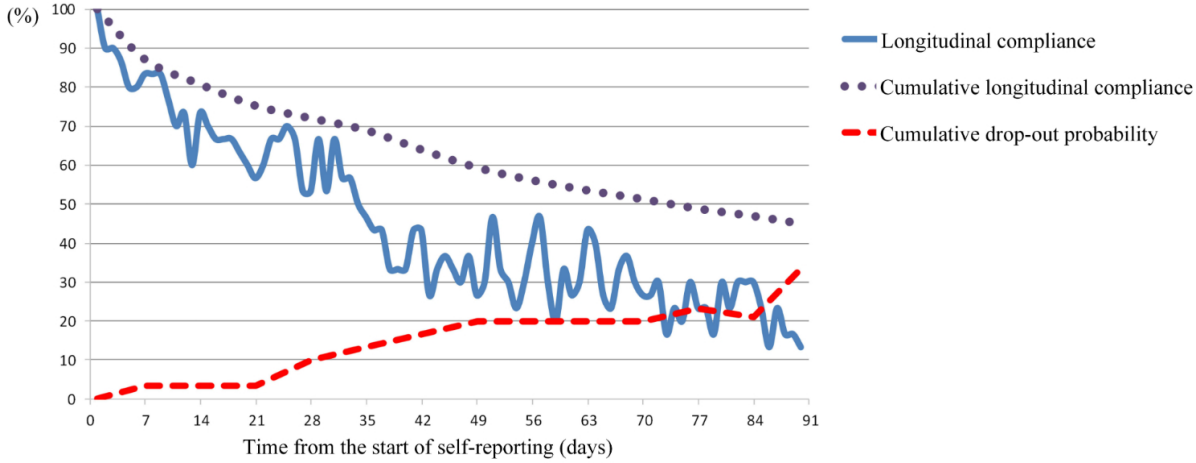


Figure 4. Changes in compliance over time.



Variables Associated With Compliance

The distribution of individual patient-level reporting rates depicted in Figure 3 exhibited a bimodal pattern, from which the median cutoff for dichotomization into higher and lower compliance groups was determined. The only variable significantly associated with greater compliance was occupational status, such that women without any kind of employment were associated with a higher rate of compliance (*P*=.03; Table 1). Furthermore, the individual patient-level reporting rate of this jobless subgroup was significantly higher than that of those with some form of employment (55.9%, SD 25.7 and 30.7%, SD 19.2, respectively; *P*=.006; Table 2). Average individual patient-level reporting rates in each subgroup classified according to the clinicopathologic variables are summarized in Table 2. Age, educational attainment, marital status, cohabitation status, and baseline BDI and HRQOL scores were not significantly associated with compliance. This study cohort was clinically homogeneous and the lymph-node status, hormone receptor expression, and chemotherapy regimens did not differ significantly between the two compliance subgroups (Table 1).

The interval from enrollment to first self-reporting (ie, lag time) was also investigated in this study. As shown in Figure 1, a push notification that asked “Did you check your journals today?” was sent daily to the patients’ smartphones from the day after enrollment and the participants were able to begin their self-reporting via the app from that day at their leisure. There was no additional reminder process via any type of online or offline tool. The intervals from enrollment to first self-report are depicted in Figure 5 (median 1 day; range 1-16 days). The mean individual patient-level reporting rate was higher for the subgroup with a 1-day lag time, defined as starting to self-report on the day immediately after enrollment, than for those with a lag of 2 or more days (51.6%, SD 24.0 and 29.6%, SD 25.3, respectively; *P*=.03).

Table 2. Average individual patient-level reporting rates in each subgroup classified according to the clinicopathologic variables (n=30).

Subgroup classified according to each variable		n (%)	Individual patient-level reporting rate (%)		
			Mean (SD)	Range	*P*
Age					NS[a]
	Young age (≤49 years)	22 (73)	46.0 (23.8)	7.8-94.4	
	Old age (≥50 years)	8 (27)	42.4 (33.1)	6.7-95.6	
Level of educational attainment					NS
	Up to high school	16 (53)	40.3 (24.8)	6.7-95.6	
	College or greater	14 (47)	50.3 (27.3)	7.8-94.4	
Marital status					NS
	Married	26 (87)	45.6 (26.4)	6.7-95.6	
	Single	1 (3)	50.0		
	Divorced	3 (10)	37.8 (31.8)	7.8-71.1	
Cohabiting					NS
	Yes	28 (93)	43.6 (26.1)	6.7-95.6	
	No	2 (7)	65.0 (21.2)	50.0-80.0	
Occupation					.006
	Yes (of any kind)	13 (43)	30.7 (19.2)	6.7-71.1	
	No (eg, stay-at-home mother)	17 (57)	55.9 (25.7)	10.0-95.6	
Baseline anxiety status with BDI [b]					NS
	No anxiety (BDI≤15)	21 (70)	46.2 (27.8)	7.8-95.6	
	Anxiety (BDI≥16)	9 (30)	42.2 (22.7)	6.7-80.0	
Pain/discomfort status with EQ5D-3L [c]					NS
	No problems	19 (63)	41.2 (27.5)	6.7-95.6	
	Problems	11 (37)	51.6 (23.0)	24.4-80.0	
Anxiety/depression status with EQ5D-3L					NS
	No problems	14 (47)	49.6 (29.9)	7.8-95.6	
	Problems	16 (53)	41.0 (22.3)	6.7-78.9	
Disease status					NS
	Localized	7 (23)	41.6 (31.4)	13.3-95.6	
	Advanced	23 (77)	46.0 (24.9)	6.7-94.4	
Hormone receptor status					NS
	Positive	18 (60)	52.8 (24.8)	6.7-95.6	
	Negative	12 (40)	33.2 (24.2)	7.8-94.4	
Neoadjuvant chemotherapy regimen					NS
	AC[d]#4	7 (23)	41.6 (31.4)	13.3-95.6	
	AC#4 followed by docetaxel#4	19 (64)	50.7 (24.6)	6.7-94.4	
	FEC[e]#3 followed by docetaxel#3	4 (13)	23.9 (11.1)	7.8-33.3	

[a]Not significant.

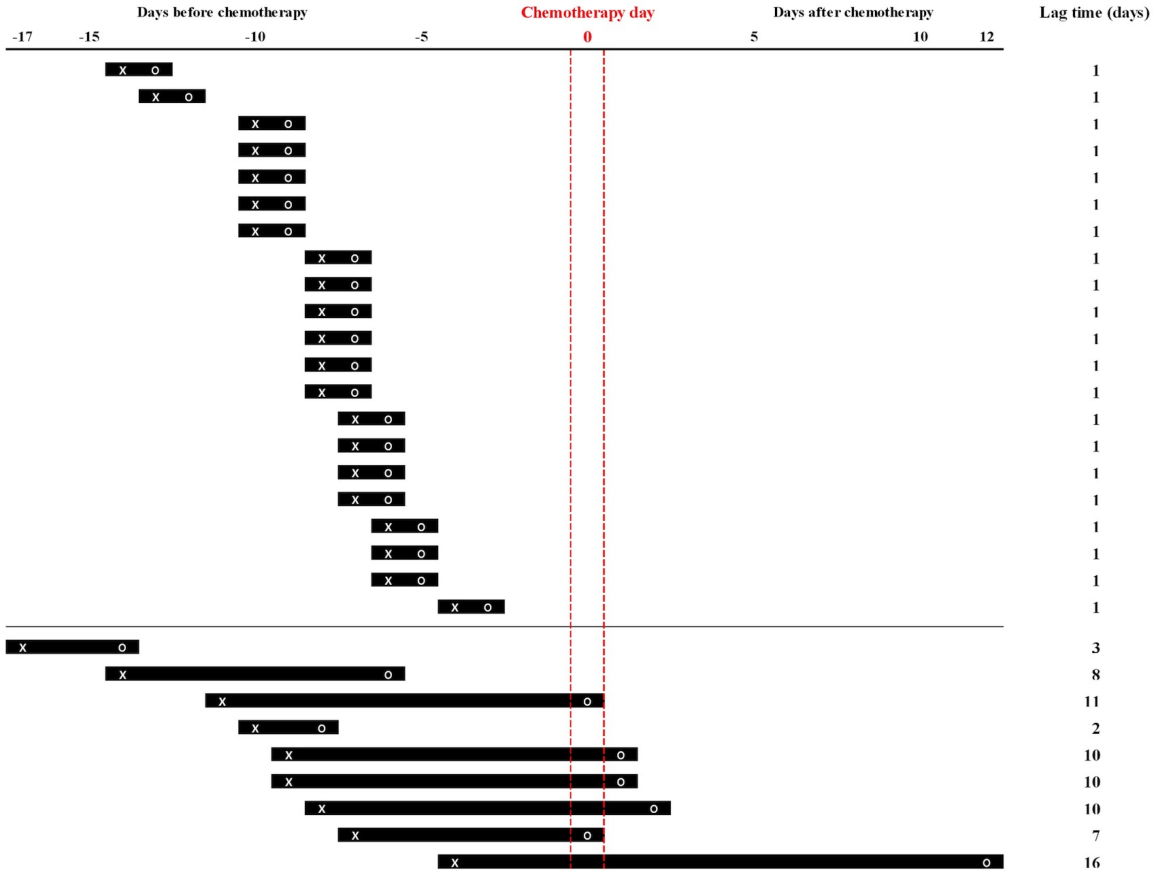
[b]BDI: Beck's Depression Inventory.

[c]EQ5D-3L: EuroQol Five Dimensional Questionnaire.

[d]Combination of doxorubicin+cyclophosphamide chemotherapy.

[e]Combination of 5-fluorouracil+epirubicin+cyclophosphamide chemotherapy.

Figure 5. Intervals from enrollment to first self-reporting. X and O indicate day of enrollment and day of start of self-reporting, respectively.



Primary Reasons for Missing Self-Reporting of Sleep Disturbance Data

A total of 24 participants responded to the question regarding their primary reason for missing self-reporting. The most common response, which accounted for 38% (9/24) of all responses, was that "The app didn't work properly." Other reasons included "I forgot" (29%, 7/24) and "I didn't think it was useful" (21%, 5/24). Minor reasons were "I was too sick" (8%, 2/24) and "I didn't feel like it" (4%, 1/24). No missing data were due to either "It was inconvenient" or "I was too busy" (Table 3). As shown in Table 3, average individual patient-level reporting rates in each subgroup categorized according to the reasons were not statistically different.

Table 3. Reasons for missing a self-reporting event and average individual patient-level reporting rates (n=24).

		Individual patient-level reporting rates (%)
Reason	n (%)	mean (SD)[a]
The app didn't work properly[b]	9 (38)	49.6 (24.6)
I forgot	7 (29)	32.4 (23.0)
I didn't think it was useful	5 (21)	57.8 (26.6)
I was too sick	2 (8)	42.2 (40.9)
I didn't feel like it	1 (4)	33.3
It was inconvenient	0 (0)	NA[c]
I was too busy	0 (0)	NA[c]
No response	6	45.0 (31.2)

[a]Means among the subgroups were not significantly different by ANOVA (analysis of variance).

[b]Temporary dysfunctions such as delay or failure of log-on or abnormal shutdown of the app during the self-reporting.

[c]Not available.

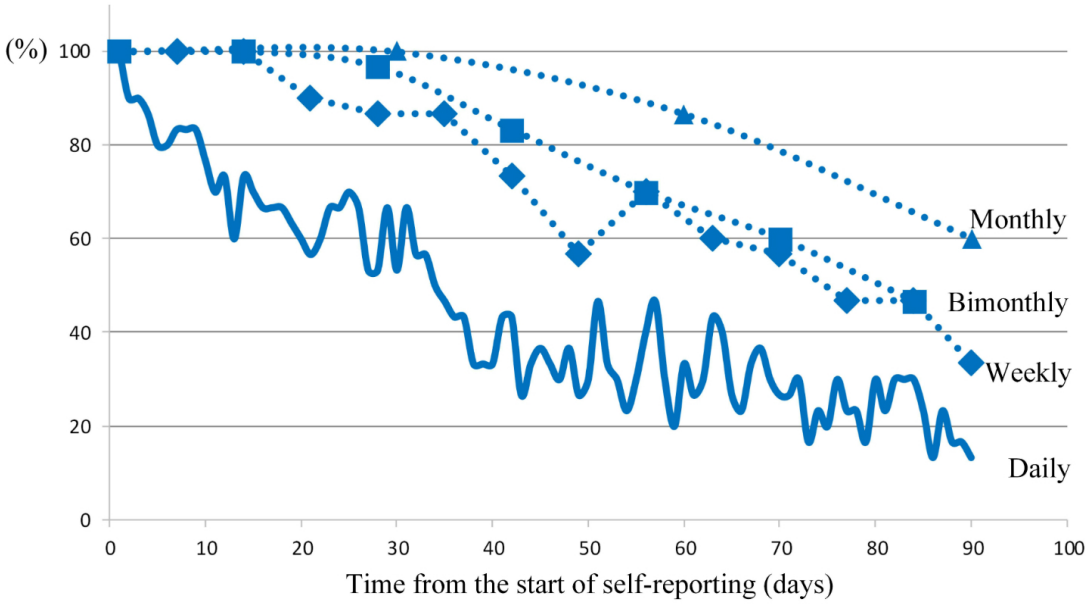
Discussion

Principal Findings

This study has shown that the collection of patient self-reporting symptoms via a smartphone at daily intervals for a relatively long-term period may be feasible in women with a recent diagnosis of breast cancer and receiving ongoing chemotherapy. The overall rate of compliance was 45.0% and the median value of individual compliance (ie, individual patient-level reporting rates) was 41.1%. About half of the participants did not self-report at around the 5th week after starting to use the app and only 13.3% of the participants could be expected to self-report at the 90th day. We found that in this population, depression, HRQOL status, and demographic characteristics such as age and educational level did not affect compliance, but the results suggest that women who were not currently in employment and those who started to use the app on the day immediately after enrollment exhibited greater compliance with daily self-reporting.

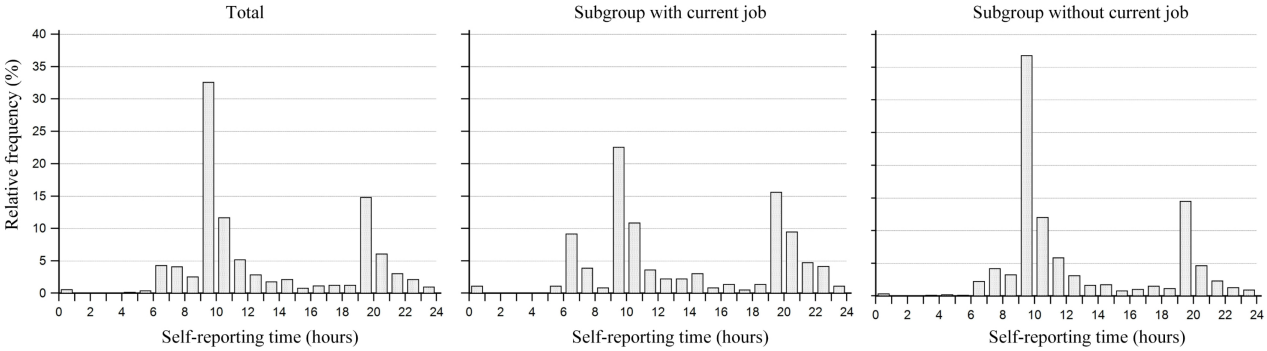
During the past decade, the greatest upsurge of mHealth research occurred between 2007 and 2008, when the new generation of smartphones, such as the iPhone and similar devices, were introduced [15]. Since most mHealth apps in mHealth research have focused on fields such as chronic conditions [15], followed by prevention/well-being and acute conditions, the evidence around smartphones with apps in cancer patients is extremely limited, particularly regarding the feasibility of nonclinic-based and daily interval data collection, as in this study. Previous studies not involving the use of smartphones found that the completion of online questionnaires on toxicity in patients receiving chemotherapy was associated with a high compliance rate (on average more than 70-80%) for self-reporting at office visits, but only 15.0% actively self-reported from home between visits [7,8]. However, our overall compliance rate was 45.0%, which is higher than found in previous studies, especially considering that the responses originated from the home between clinic visits and the response interval was daily. In addition, our overall compliance of 45.0% was calculated based on daily self-reporting and so can be considered as daily compliance. If patient-level compliance is quantified, defined as the proportion of each time unit during which a given patient self-reported at least once, as in one previous study [9], our weekly, bimonthly, and monthly compliance rates increase to 69.7%, 76.1%, and 82.2%, respectively, all of which are above a reasonable cutoff value by experts (Figure 6) [9]. Previous research indicates that the main barrier to self-reporting via the Internet is not receiving a reminder while at home (ie, participants simply forget) [8,9,14]. We can assume that the regular push notifications available with this kind of app may be responsible for the higher-than-expected compliance rate seen in the present study.

Figure 6. Comparison of longitudinal compliance rates according to different time units.



The lower compliance observed among women with a current job may be due to failure to immediately respond to push notifications and simple forgetfulness. Self-reporting during the hour immediately following the push notification (9-10 am and 7-8 pm) accounted for almost half of all responses, which is consistent with the proportion of women in this cohort without a current job. However, in women with current employment, the pattern of self-reporting had two wider peaks (Figure 7). To improve the rate of an immediate response to push notification, various schedules for push notification timing from which patients can choose should be developed into the app in future studies. We cautiously expect that such tailored push notifications will improve compliance in the subgroup of the currently employed, since, as suggested by the data in Table 3, simple forgetfulness rather than being busy was the primary reason for missing self-reporting.

Figure 7. Push notifications and distribution of self-reporting time.



The compliance rate was higher for the subgroup without a lag time in this study (mean 51.6%, SD 24.0 and 29.6%, SD 25.3, respectively; *P*=.03). One possible explanation for this is that the interval between enrollment to the first self-reporting event may be a surrogate variable for self-efficacy, which is defined as the extent or strength of one's belief in one's own ability to complete tasks and reach goals [16]. The concept of self-efficacy has been receiving increasing recognition as a predictor of changes and maintenance of health behavior [17], as well as of the psychological attitudes of cancer patients [18]. A general self-efficacy questionnaire was not administered in this study and so we cannot draw any definite conclusions regarding self-efficacy and compliance. Further investigation focusing on self-efficacy rather than conventional patients' characteristics such as age, educational level, marital status, cohabitation status, and mood—which were not significantly associated with self-reporting compliance in this study—might reveal practical and modifiable predictors for selecting the subgroup with the highest compliance rate in the emerging smartphone-based health care field.

In this feasibility study, we also sought to determine the degree of persistence of compliance and the optimal time interval for self-reporting of daily changing symptoms, such as the sleep-disturbance data found in this study. Although it is not possible to form categorical conclusions regarding these parameters from the findings of the present single study, the results will help toward the development of future strategies for collecting repetitive biodata. Researchers working on smartphone-based health care can expect interactive and responsive communications from 50% of women receiving chemotherapy at the 5th week, from 20-50% between the 6th and 10th week and from just 10-30% thereafter (Figure 4). In addition, given daily collection of self-reporting data via smartphone with an overall compliance of 50%, the entire collection period should be less than 10 weeks, corresponding to a cumulative longitudinal day-level reporting rate of 51.3% (Figure 4). The rationale regarding the dropout cutoff in the Methods section was that we considered it unreasonable to keep encouraging daily self-reporting of symptoms by those who respond on average almost weekly or less. From the cumulative dropout probability curve, we predicted that 10%, 20%, and 21% of participants would drop out at days 28, 56, and 84, respectively (Figure 4).

Limitations

The limitations of this study include the enrollment only of Korean women with breast cancer at a single urban tertiary cancer center. Other populations may exhibit different self-reporting characteristics in terms of compliance and should be independently evaluated. In addition, these findings should be generalized only after taking into account that this study was conducted under special circumstances in South Korea in which there is easy and ubiquitous access to wireless networks and the penetration of smartphones is the second highest in the world, after the United Arab Emirates (Multimedia Appendix 1). The study design meant that only patients possessing a smartphone with app avidity were included. Of the total population of patients who were approached regarding participation in this study, 13% (9/67) were excluded by this criterion. Therefore, the results of the present study should be interpreted with caution: the compliance rates in this population may have been overestimated. Nonetheless, this feasibility study provides an initial understanding of the opportunities for successful smartphone-based collection of real-time, self-reporting data, and valuable insights into the development of more practicable interventions with smartphones in the real cancer care setting.

Conclusions

In conclusion, the findings of this study suggest that 90-day, longitudinal collection of daily self-reporting sleep-disturbance data via a smartphone app is feasible. Further research is needed to determine how to sustain the compliance with a self-reporting program over a longer period of time and to select subpopulations with higher compliance rates for mobile health care.

Acknowledgments

This study was supported by a grant from the National R&D Program for Cancer Control, Ministry for Health and Welfare, Republic of Korea (no. 1220170).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Top 15 countries with highest smartphone penetration in 2013 (from our mobile planet by Google).

[JPG File, 100KB - jmir_v16i5e135_app1.jpg]

References

1. Consulting VW. UN Foundation-Vodafone Foundation Partnership. 2009. mHealth for development: the opportunity of mobile technology for healthcare in the developing world URL: http://www.unfoundation.org/news-and-media/publications-and-speeches/mhealth-for-development-1.html [accessed 2014-05-21] [WebCite Cache ID 6Pk02lsIc]
2. Ryan D, Cobern W, Wheeler J, Price D, Tarassenko L. Mobile phone technology in the management of asthma. J Telemed Telecare 2005;11 Suppl 1:43-46. [doi: 10.1258/1357633054461714] [Medline: 16035991]
3. Anhøj J, Møldrup C. Feasibility of collecting diary data from asthma patients through mobile phones and SMS (short message service): response rate analysis and focus group evaluation from a pilot study. J Med Internet Res 2004;6(4):e42. [doi: 10.2196/jmir.6.4.e42] [Medline: 15631966]
4. Weaver A, Young AM, Rowntree J, Townsend N, Pearson S, Smith J, et al. Application of mobile phone technology for managing chemotherapy-associated side-effects. Ann Oncol 2007 Nov;18(11):1887-1892 [FREE Full text] [doi: 10.1093/annonc/mdm354] [Medline: 17921245]
5. Killingsworth MA, Gilbert DT. A wandering mind is an unhappy mind. Science 2010 Nov 12;330(6006):932 [FREE Full text] [doi: 10.1126/science.1192439] [Medline: 21071660]
6. Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient non-compliance with paper diaries. BMJ 2002 May 18;324(7347):1193-1194 [FREE Full text] [doi: 10.1136/bmj.324.7347.1193] [Medline: 12016186]
7. Basch E, Artz D, Dulko D, Scher K, Sabbatini P, Hensley M, et al. Patient online self-reporting of toxicity symptoms during chemotherapy. J Clin Oncol 2005 May 20;23(15):3552-3561. [doi: 10.1200/JCO.2005.04.275] [Medline: 15908666]
8. Basch E, Iasonos A, Barz A, Culkin A, Kris MG, Artz D, et al. Long-term toxicity monitoring via electronic patient-reported outcomes in patients receiving chemotherapy. J Clin Oncol 2007 Dec 1;25(34):5374-5380. [doi: 10.1200/JCO.2007.11.2243] [Medline: 18048818]
9. Judson TJ, Bennett AV, Rogak LJ, Sit L, Barz A, Kris MG, et al. Feasibility of long-term patient self-reporting of toxicities from home via the Internet during routine chemotherapy. J Clin Oncol 2013 Jul 10;31(20):2580-2585. [doi: 10.1200/JCO.2012.47.6804] [Medline: 23733753]
10. de Jongh T, Gurol-Urganci I, Vodopivec-Jamsek V, Car J, Atun R. Mobile phone messaging for facilitating self-management of long-term illnesses. Cochrane Database Syst Rev 2012;12:1-52. [doi: 10.1002/14651858.CD007459.pub2] [Medline: 23235644]
11. Gentry S, van-Velthoven MH, Tudor Car L, Car J. Telephone delivered interventions for reducing morbidity and mortality in people with HIV infection. Cochrane Database Syst Rev 2013;5:1-68. [doi: 10.1002/14651858.CD009189.pub2] [Medline: 23728687]
12. Whittaker R, McRobbie H, Bullen C, Borland R, Rodgers A, Gu Y. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2012;11:1-24. [doi: 10.1002/14651858.CD006611.pub3] [Medline: 23152238]
13. Lee G, Lee JW, Park JY, Shin YW, Jung KH, Oh EY, et al. "Pit-A-Pat": Development of a Mobile Health Application for Breast Cancer Patients. 2012 Presented at: 2012 mHealth Summit; Dec 3-5, 2012; Washington D.C., USA p. 3-5.
14. Basch E, Artz D, Iasonos A, Speakman J, Shannon K, Lin K, et al. Evaluation of an online platform for cancer patient self-reporting of chemotherapy toxicities. J Am Med Inform Assoc 2007;14(3):264-268 [FREE Full text] [doi: 10.1197/jamia.M2177] [Medline: 17329732]
15. Fiordelli M, Diviani N, Schulz PJ. Mapping mHealth research: a decade of evolution. J Med Internet Res 2013;15(5):e95 [FREE Full text] [doi: 10.2196/jmir.2430] [Medline: 23697600]
16. Ormrod JE. Educational Psychology: Developing Learners. Upper Saddle River, N. Upper Saddle River, N.J: Pearson/Merrill Prentice Hall; 2008.
17. Strecher VJ, DeVellis BM, Becker MH, Rosenstock IM. The role of self-efficacy in achieving health behavior change. Health Educ Q 1986;13(1):73-92. [Medline: 3957687]
18. Cunningham AJ, Lockwood GA, Cunningham JA. A relationship between perceived self-efficacy and quality of life in cancer patients. Patient Educ Couns 1991 Feb;17(1):71-78. [doi: 10.1016/0738-3991(91)90052-7] [Medline: 1997998]

Abbreviations

app: application
BDI: Beck's Depression Inventory
eHealth: electronic health

ePROs: electronic PROs
EQ-5D-3L: EuroQol Five Dimensional Questionnaire
HRQOL: health-related quality of life
ICT: information and communication technology
mHealth: mobile health
OS: operating system
PRO: patient-reported outcome

Edited by G Eysenbach; submitted 25.03.14; peer-reviewed by P Dykes, J Hopper; comments to author 12.04.14; revised version received 14.04.14; accepted 19.05.14; published 23.05.14.

Please cite as:
Min YH, Lee JW, Shin YW, Jo MW, Sohn G, Lee JH, Lee G, Jung KH, Sung J, Ko BS, Yu JH, Kim HJ, Son BH, Ahn SH
Daily Collection of Self-Reporting Sleep Disturbance Data via a Smartphone App in Breast Cancer Patients Receiving Chemotherapy: A Feasibility Study
J Med Internet Res 2014;16(5):e135
URL: http://www.jmir.org/2014/5/e135/
doi:10.2196/jmir.3421
PMID:24860070

Original Paper

Do Cancer Patients Tweet? Examining the Twitter Use of Cancer Patients in Japan

Atsushi Tsuya[1], PhD; Yuya Sugawara[2], MS; Atsushi Tanaka[1], PhD; Hiroto Narimatsu[3], MD, PhD

[1]Yamagata University Graduate School of Science and Engineering, Yamagata University, Yonezawa, Japan

[2]Medical Informatics focus, Environmental Life Science major, Yamagata University Graduate School of Medical Science, Yamagata University, Yamagata, Japan

[3]Department of Public Health, Yamagata University Graduate School of Medicine, Yamagata University, Yamagata, Japan

Corresponding Author:
Hiroto Narimatsu, MD, PhD
Department of Public Health
Yamagata University Graduate School of Medicine
Yamagata University
2-2-2
Iida-nishi
Yamagata, 990-9585
Japan
Phone: 81 (0)23 628 5260
Fax: 81 (0)23 628 5261
Email: hiroto-narimatsu@umin.org

Abstract

Background: Twitter is an interactive, real-time media that could prove useful in health care. Tweets from cancer patients could offer insight into the needs of cancer patients.

Objective: The objective of this study was to understand cancer patients' social media usage and gain insight into patient needs.

Methods: A search was conducted of every publicly available user profile on Twitter in Japan for references to the following: breast cancer, leukemia, colon cancer, rectal cancer, colorectal cancer, uterine cancer, cervical cancer, stomach cancer, lung cancer, and ovarian cancer. We then used an application programming interface and a data mining method to conduct a detailed analysis of the tweets from cancer patients.

Results: Twitter user profiles included references to breast cancer (n=313), leukemia (n=158), uterine or cervical cancer (n=134), lung cancer (n=87), colon cancer (n=64), and stomach cancer (n=44). A co-occurrence network is seen for all of these cancers, and each cancer has a unique network conformation. Keywords included words about diagnosis, symptoms, and treatments for almost all cancers. Words related to social activities were extracted for breast cancer. Words related to vaccination and support from public insurance were extracted for uterine or cervical cancer.

Conclusions: This study demonstrates that cancer patients share information about their underlying disease, including diagnosis, symptoms, and treatments, via Twitter. This information could prove useful to health care providers.

(J Med Internet Res 2014;16(5):e137) doi:10.2196/jmir.3298

KEYWORDS

communication; co-occurrence; Internet; leukemia; Web 2.0

Introduction

Twitter: A Novel Social Media

Twitter is a free social networking and micro-blogging service that enables its millions of users to send and read each other's "tweets", or short messages limited to 140 characters. The users themselves determine whether their tweets can be read by the general public or should be restricted to preselected "followers". As of March 2012, the service had more than 200 million registered users and processed about 400 million tweets per day [1,2].

A recent analysis of the "Twitter stream" revealed that a substantial proportion of tweets contain general chatter, that is, user-to-user conversations that are of interest only to the parties

involved, links to interesting pieces of news, or spam and self-promotion [1]. Despite the high level of noise, the Twitter stream does contain useful information. Recently, we and other researchers demonstrated that Twitter is emerging as an important channel for communicating about cancer [3-7]. Many recent news events or scientific issues have been documented and discussed via Twitter directly from users on the site in real time [8]. Although the information that one tweet includes is limited, Twitter can convey more immediacy with interactivity than website homepages or blogs [1,9-12], such as the Association of Cancer Online Resources [9]. Thus, Twitter has the potential to play a different role in sharing medical information among patients.

Twitter in Cancer Patients

In a recent case study, we demonstrated that Twitter networks of cancer patients centered on active users and that these networks could provide psychological support for cancer patients [4]. Because of certain restrictions of the search tool, the study was not able to conduct a large-scale comprehensive qualitative analysis. Therefore, in the present study, we examine cancer patients' social media usage by analyzing the data with a text mining method using an application programming interface (API) [2]. Thus, we were able to comprehensively analyze the Twitter data of cancer patients on a large scale.

Methods

Search for Twitter Accounts of Cancer Patients

A search was conducted of every publicly available user profile on Twitter in Japan. We examined the number of user accounts in which the names of cancers are described in the profile. The search terms included breast cancer, leukemia, colon cancer, rectal cancer, colorectal cancer, uterine cancer, cervical cancer, stomach cancer, lung cancer, and ovarian cancer. These names were alternatively searched using "cancer" in the Japanese hiragana and katakana writing system and in Chinese characters. The site used for the profile search was "16 (one-six) Profile Search β Version for Twitter" [13], which enabled us to search, in addition to profiles, the number of follows, followers, tweets, lists, registered dates, and last posted dates. The search was conducted on August 18, 2013. This study was approved by the Institutional Review Board at Yamagata University Faculty of Medicine (H24-133).

Content Analysis of Tweets

Using Twitter API, the latest tweets (maximum 200 tweets) from each account, found after the above search, were gathered. Twitter API is a function officially provided by the organization that operates Twitter to Twitter application developers in order to provide useful and convenient functions to Twitter users. By incorporating Twitter API into an application, the application developer can add Twitter functions such as Twitter search results or obtaining tweets from Twitter accounts [14].

First, tweets obtained from each account through Twitter API were separated onto different lines with a period ".". Subsequently, these were broken down into morphemes ("words") using the Japanese language morpheme analysis software ChaSen (from the Nara Institute of Science and Technology, Japan). Here, the words were represented in their original forms. Nouns were then extracted from these words and were listed on separate lines. These nouns ("noun group") listed in separate lines were then grouped together by account. Occasionally, verbs and adjectives are also extracted with text mining. However, in the present study, we did not extract verbs and adjectives for the following reasons: (1) difficulties in dealing with negative sentences, and (2) low percentage of the part of speech of the extracted word. In addition, nouns obtained that were synonyms were integrated into one noun. Synonyms were determined by the authors by referring to WordNet Web search services [15]. Dictionaries that contained words obtained from the descriptions on websites were used as the default for ChaSen ("cancer information services" [16] and "good health care" [17]).

Tweets were obtained during the following dates and times: 0:39–2:52 on August 19, 2012, for stomach cancer, colon and colorectal cancer, and leukemia tweets; 14:40–17:24 on August 20, 2012, for uterine cancer, breast cancer, and lung cancer tweets.

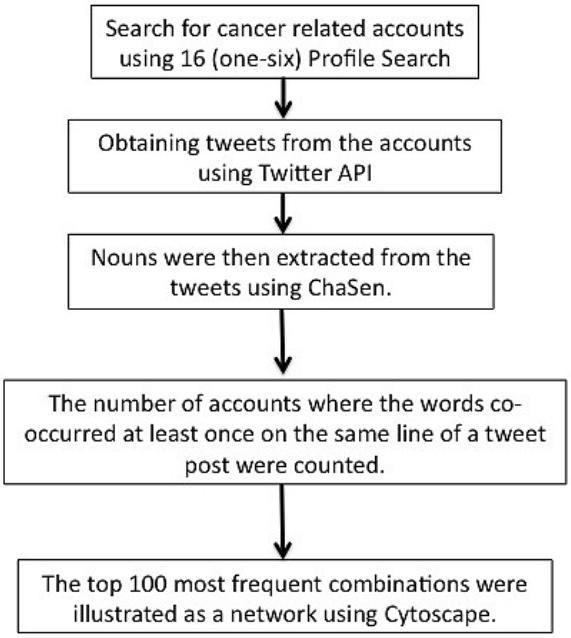
Generation of Co-Occurrence Networks

The procedure of generating the co-occurrence network is shown in Figure 1. Co-occurrence is the relation between the keywords that appear together in each tweet; thus, co-occurrence means a close relationship between words. In this study, we demonstrate the features of tweets by cancer patients by analyzing the co-occurrence of keywords.

To accomplish this, we created co-occurrence networks using the following procedure: (1) the tweets from the cancer-related accounts were broken down into words using ChaSen, (2) from the noun groups that were combinations of two words, we counted the number of accounts where the words co-occurred at least once on the same line of a tweet, and (3) from the word combinations that co-occurred on the same line of a tweet, the top 100 most frequent combinations (the top 100 in number of accounts) were illustrated as a network with words depicted as nodes and combinations as links. Network analysis software Cytoscape [18] was used for the illustration. We first used the spring model as a node placement rule and subsequently made adjustments such that each word and each link overlapped as little as possible. The spring model is a method that can illustrate networks from the perspective of evenness of side length as well as uniformity and symmetry of node distribution. It regards each side as a spring that follows Hooke's law and each node as an electrically charged particle that follows Coulomb's law, and the layout is established by determining the equilibrium state [19].

In the method we used to create co-occurrence networks in this study, as a way to handle the high frequency of extremely specialized tweets, the co-occurrence frequency of co-occurrence networks was defined as the number of accounts where words co-occurred in tweets, rather than the number of co-occurrences of words, which is typically done when creating co-occurrence networks. This then prevented extremely specialized words completely unrelated to cancer from appearing in the co-occurrence networks.

Figure 1. Procedure for generating the co-occurrence network.



Results

The accounts we searched included references to breast cancer (n=313), leukemia (n=158), uterine and cervical cancer (n=134), lung cancer (n=87), colon cancer (n=64), and stomach cancer (n=44). The co-occurrence networks of those cancers are shown in Figure 2-7. Table 1 summarizes the keywords from tweets related to different types of cancer. Each cancer has a unique network conformation. The keywords included words about diagnosis, symptoms, and treatments for almost all cancers.

Table 1. Summary of keywords in tweets according to cancer type.

	Diagnosis	Symptoms	Treatments	Others
Stomach cancer	CT[a], MRI[b], tumor marker	Lumbago, TS-1, side effects	Anti-cancer drug, TS-1, administration of iron	Not available
Colon and colorectal cancer	CT, PET[c]	ELPLAT, side effects	Chemotherapy, diet	Nursing care
Cancer of uterus and cervical cancer	Not available	Lymphedema	Not available	Educational activity, screening, not covered by health insurance, vaccination, official support
Lung cancer	CT	Metastasis, shoulder pain, back pain, Iressa, side effects	Anti-cancer drug, Iressa, Tarceva	Palliative care
Breast cancer	Self-diagnosis	Metastasis, lymphedema	Chemotherapy, hormonal treatment	Palliative care, the pink ribbon
Leukemia	Liver function test	Liver function test, foot pain, immunosuppression, GVHD[d]	Chemotherapy, steroid treatment, transfusion of red blood cells, platelet transfusion	AML[e], hematopoietic stem cell transplantation

[a]CT: computed tomography.

[b]MRI: magnetic resonance imaging.

[c]PET: positron emission tomography.

[d]GVHD: graft-versus-host disease.

[e]AML: acute myeloid leukemia.

Figure 2. Co-occurrence network of cancers: breast cancer.

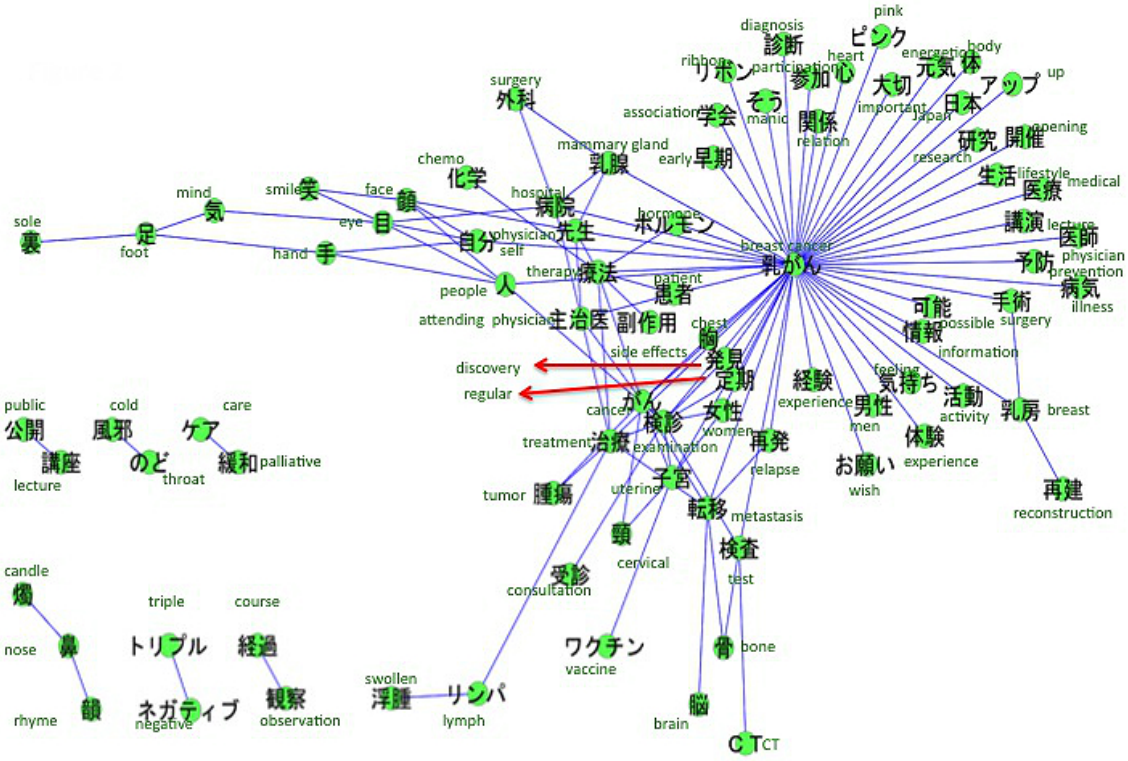


Figure 3. Co-occurrence network of cancers: leukemia.

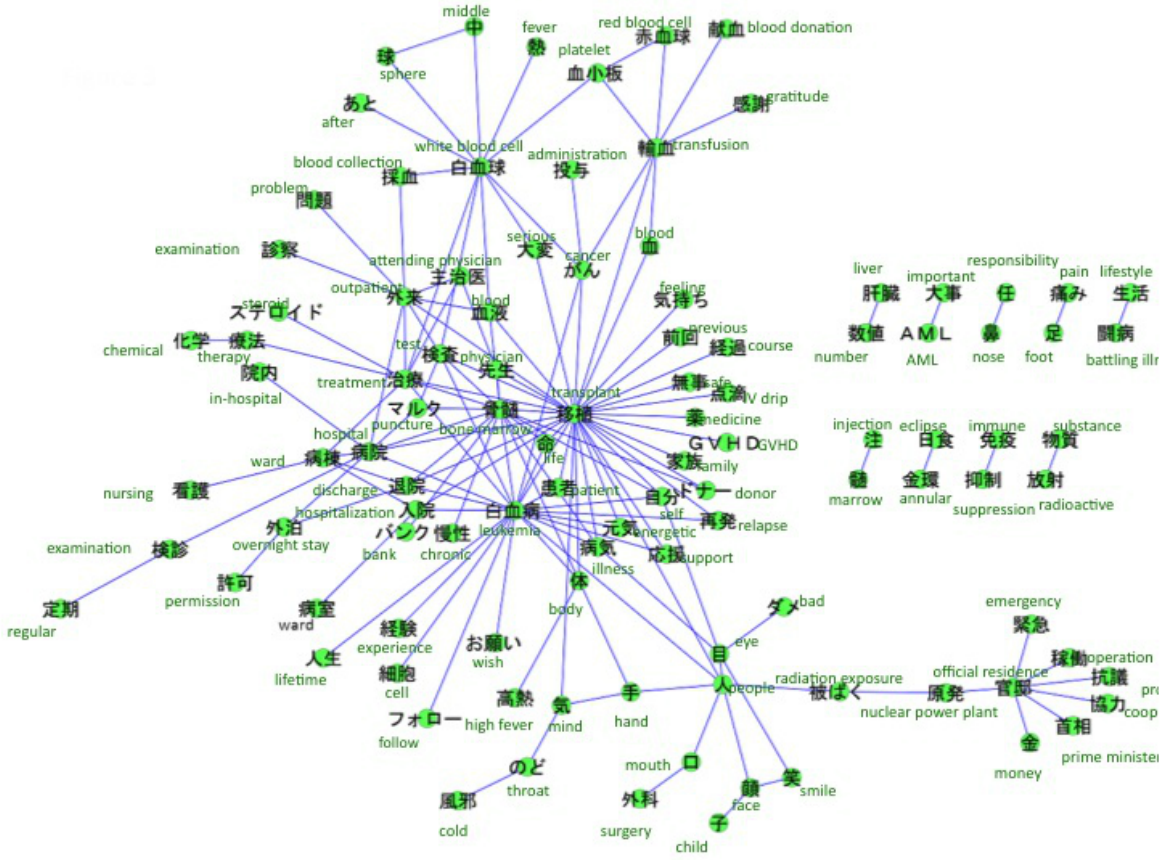


Figure 4. Co-occurrence network of cancers: uterine and cervical cancer.

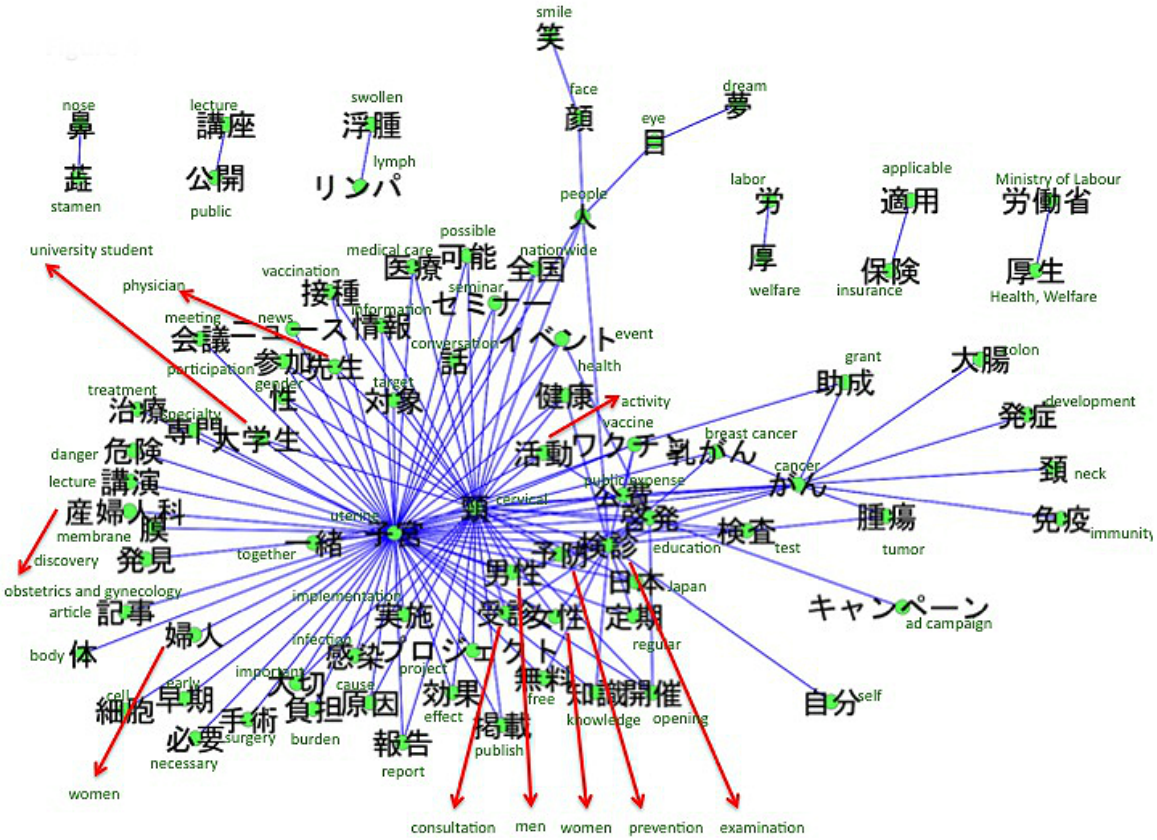


Figure 5. Co-occurrence network of cancers: lung cancer.

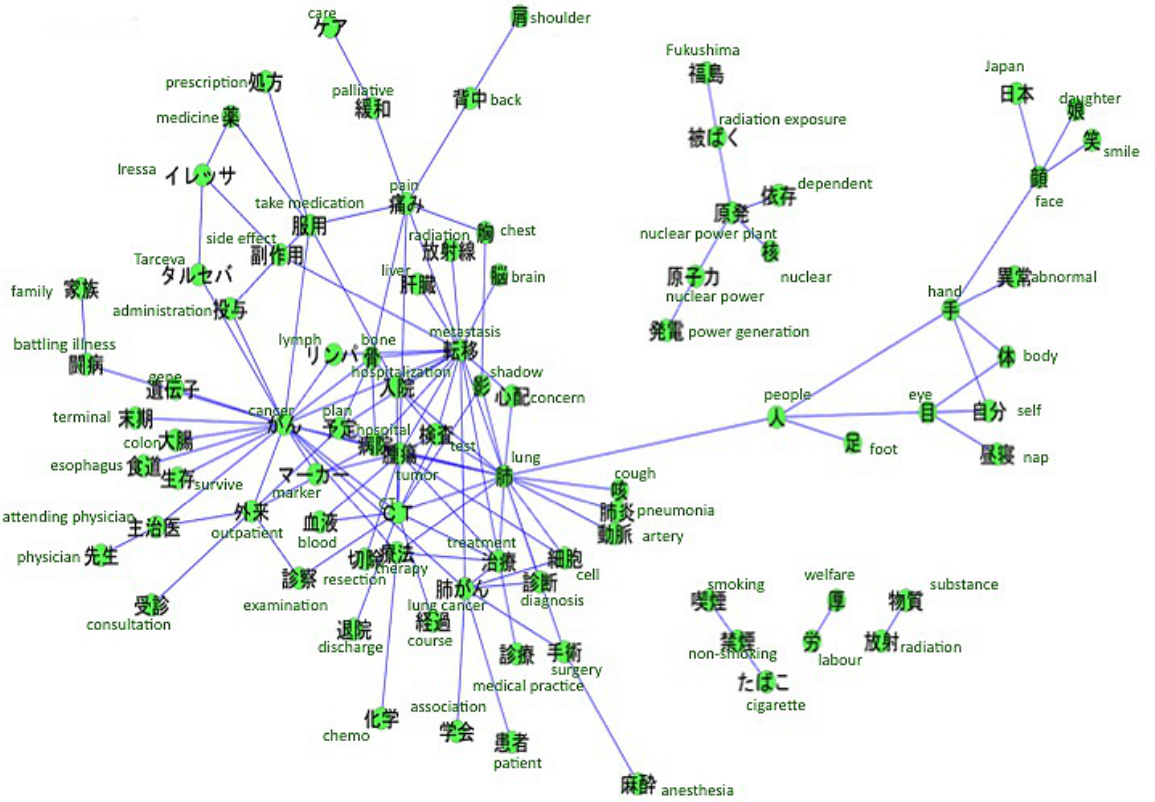


Figure 6. Co-occurrence network of cancers: stomach cancer.

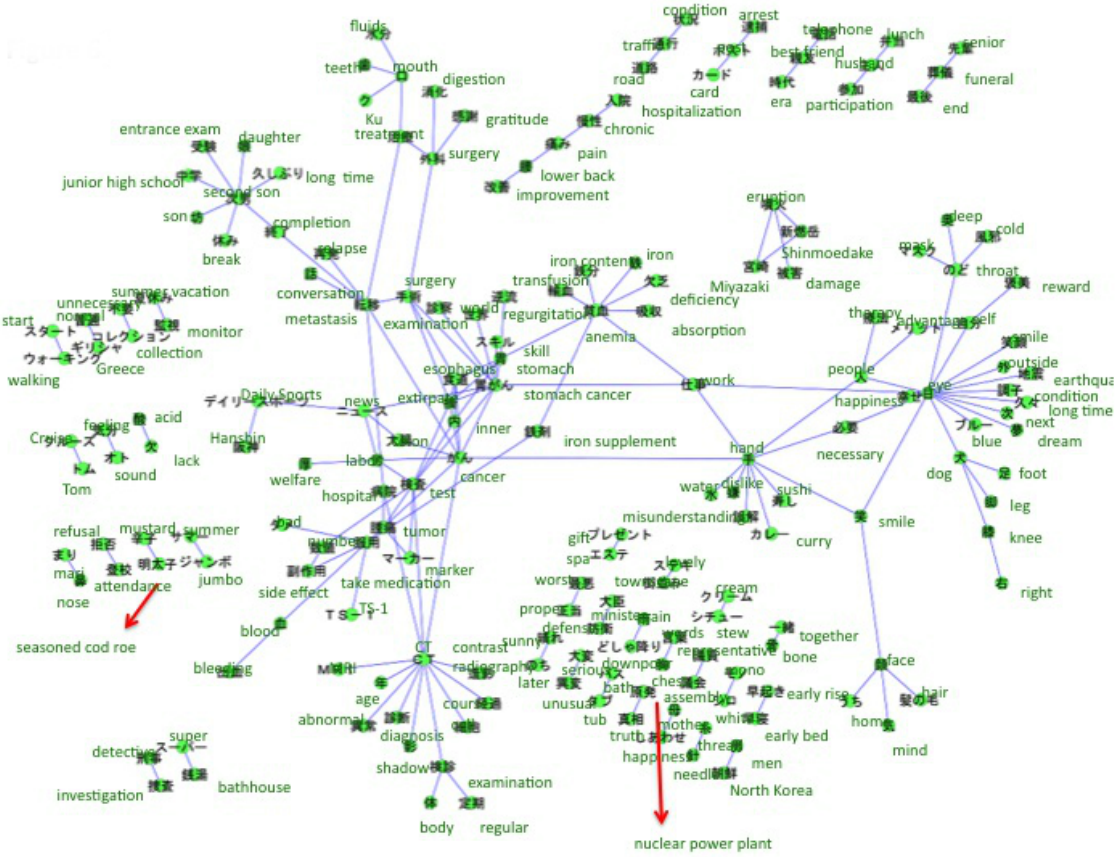
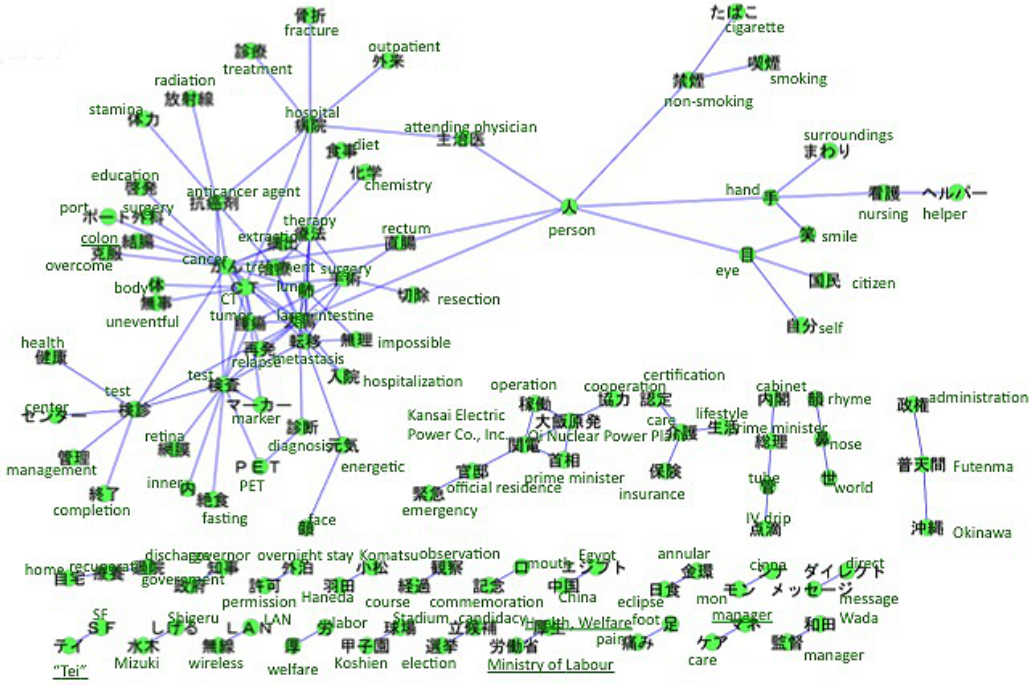


Figure 7. Co-occurrence network of cancers: colon and colorectal cancer.



Discussion

Comprehensive Analysis of Tweets

In this study, we used an information technology procedure to comprehensively analyze the content of cancer patients' tweets. In previous studies, researchers verified each individual tweet, but this method restricted the range of Twitter information that could be obtained [4]. Moreover, a notable point of this analysis method was that we were able to exclude tweets unrelated to the diseases of interest. Using our method, information on tweets related to specific diseases can now be collected efficiently.

Although we used this method to evaluate tweets from cancer patients, in the future, we plan to apply this method to the study of other diseases, for example, lifestyle-related diseases.

Twitter data can be obtained from a variety of sources. In this study, we used Twitter API because it uses an automated approach to data retrieval and is free of charge. However, the number of tweets retrieved through Twitter API is capped at approximately 1% of all tweets, with no assurance of a random or representative sample [2]. Thus, retrieving Twitter's full data stream through automated dashboard vendors or a Twitter data reseller may provide further findings.

Tweets Related to the Cancers

This study found that information related to cancer, such as treatment, diagnosis, and symptoms, is shared among cancer patients on Twitter (Table 1). Furthermore, the extracted keywords were considered to be medically important for that specific disease, reflecting the fact that cancer patients use Twitter as a tool for sharing medical information. Additionally, depending on the type of cancer, it was clear that there were specific characteristics to the tweet content. For example, in uterine or cervical cancer and breast cancer, there were keywords not related to immediate medical care, for example, "cervical cancer vaccine" for uterine or cervical cancer and "pink ribbon" for breast cancer. These most likely indicate that patients are also affected by the heightened social interest in a cervical cancer vaccine [20] and the social excitement of the pink ribbon movement. These topics were also covered by regular news media, such as TV or newspaper. This indicates that the content of tweets can be affected by those media.

Conclusions and Future Directions

We indicated in a previous study [4] that Twitter is useful for cancer patients to exchange ordinary information. As industries obtain and utilize tweet information from Twitter as marketing tools, health care will be able to retrieve, study, and make use of tweet information. In this study, we comprehensively and efficiently collected tweet information related to diseases, demonstrating that information about cancer patients can be collected on social media. Effective use of this information will be helpful in developing cancer care that better suits the patients' needs. For example, health care providers can more effectively give information or medical services to patients, resulting in an increase in patient satisfaction.

Acknowledgments

This work was supported by a Research Grant (H24-26, 24200701 to Dr Narimatsu) from the Ministry of Health, Labour and Welfare of Japan.

Conflicts of Interest

None declared.

References

1. Signorini A, Segre AM, Polgreen PM. The use of Twitter to track levels of disease activity and public concern in the U.S. during the influenza A H1N1 pandemic. PLoS One 2011;6(5):e19467 [FREE Full text] [doi: 10.1371/journal.pone.0019467] [Medline: 21573238]
2. Kim AE, Hansen HM, Murphy J, Richards AK, Duke J, Allen JA. Methodological considerations in analyzing Twitter data. J Natl Cancer Inst Monogr 2013 Dec;2013(47):140-146. [doi: 10.1093/jncimonographs/lgt026] [Medline: 24395983]
3. De la Torre-Díez I, Díaz-Pernas FJ, Antón-Rodríguez M. A content analysis of chronic diseases social groups on Facebook and Twitter. Telemed J E Health 2012;18(6):404-408. [doi: 10.1089/tmj.2011.0227] [Medline: 22650380]
4. Sugawara Y, Narimatsu H, Hozawa A, Shao L, Otani K, Fukao A. Cancer patients on Twitter: a novel patient community on social media. BMC Res Notes 2012;5:699 [FREE Full text] [doi: 10.1186/1756-0500-5-699] [Medline: 23270426]
5. Keim-Malpass J, Steeves RH. Talking with death at a diner: young women's online narratives of cancer. Oncol Nurs Forum 2012 Jul;39(4):373-8, 406. [doi: 10.1188/12.ONF.373-378] [Medline: 22750895]
6. Lyles CR, López A, Pasick R, Sarkar U. "5 mins of uncomfyness is better than dealing with cancer 4 a lifetime": an exploratory qualitative analysis of cervical and breast cancer screening dialogue on Twitter. J Cancer Educ 2013 Mar;28(1):127-133. [doi: 10.1007/s13187-012-0432-2] [Medline: 23132231]
7. Grajales FJ, Sheps S, Ho K, Novak-Lauscher H, Eysenbach G. Social media: a review and tutorial of applications in medicine and health care. J Med Internet Res 2014;16(2):e13 [FREE Full text] [doi: 10.2196/jmir.2912] [Medline: 24518354]
8. Mandavilli A. Peer review: Trial by Twitter. Nature 2011 Jan 20;469(7330):286-287. [doi: 10.1038/469286a] [Medline: 21248816]
9. Association of Cancer Online Resources. URL: http://www.acor.org/ [accessed 2014-04-02] [WebCite Cache ID 6OWutx4Vd]
10. Narimatsu H, Matsumura T, Morita T, Kishi Y, Yuji K, Kami M, et al. Detailed analysis of visitors to cancer-related web sites. J Clin Oncol 2008 Sep 1;26(25):4219-4223. [doi: 10.1200/JCO.2008.18.7468] [Medline: 18757342]
11. Morita T, Narimatsu H, Matsumura T, Kodama Y, Hori A, Kishi Y, et al. A study of cancer information for cancer patients on the internet. Int J Clin Oncol 2007 Dec;12(6):440-447. [doi: 10.1007/s10147-007-0707-5] [Medline: 18071863]

12. Huerta TR, Hefner JL, Ford EW, McAlearney AS, Menachemi N. Hospital website rankings in the United States: expanding benchmarks and standards for effective consumer engagement. J Med Internet Res 2014;16(2):e64 [FREE Full text] [doi: 10.2196/jmir.3054] [Medline: 24568892]
13. 16 (one-six) Profile Search β Version for twitter (in Japanese). URL: http://www.16ps.jp/ [accessed 2014-02-04] [WebCite Cache ID 6N7hZgs2e]
14. Twitter API. URL: https://dev.twitter.com/docs/api/1.1 [accessed 2014-05-23] [WebCite Cache ID 6PmkuCscu]
15. Weblio. URL: http://thesaurus.weblio.jp/ [accessed 2014-05-23] [WebCite Cache ID 6Pml1rdPj]
16. ganjoho. URL: http://ganjoho.jp [accessed 2014-05-23] [WebCite Cache ID 6PmlAHvaZ]
17. Health Goo. URL: http://health.goo.ne.jp/ [accessed 2014-05-23] [WebCite Cache ID 6PmlHDS3H]
18. Saito R, Smoot ME, Ono K, Ruscheinski J, Wang PL, Lotia S, et al. A travel guide to Cytoscape plugins. Nat Methods 2012 Nov;9(11):1069-1076 [FREE Full text] [doi: 10.1038/nmeth.2212] [Medline: 23132118]
19. Kamada T, Kawai S. An algorithm for drawing general undirected graphs. Information Processing Letters 1989 Apr;31(1):7-15. [doi: 10.1016/0020-0190(89)90102-6]
20. Gilmour S, Kanda M, Kusumi E, Tanimoto T, Kami M, Shibuya K. HPV vaccination programme in Japan. Lancet 2013 Aug 31;382(9894):768. [doi: 10.1016/S0140-6736(13)61831-0] [Medline: 23993189]

Abbreviations

API: application programming interface

Edited by G Eysenbach; submitted 03.02.14; peer-reviewed by Y Nakata, B Rimer; comments to author 14.03.14; revised version received 03.04.14; accepted 13.04.14; published 27.05.14.

Please cite as:
Tsuya A, Sugawara Y, Tanaka A, Narimatsu H
Do Cancer Patients Tweet? Examining the Twitter Use of Cancer Patients in Japan
J Med Internet Res 2014;16(5):e137
URL: http://www.jmir.org/2014/5/e137/
doi:10.2196/jmir.3298
PMID:24867458

Original Paper

Web-Based Cognitive Training: Patient Adherence and Intensity of Treatment in an Outpatient Memory Clinic

Vítor Tedim Cruz[1,2], MD; Joana Pais[1], PhD; Ivânia Alves[1], MD; Luís Ruano[1], MD; Cátia Mateus[1], MSc; Rui Barreto[1], MD; Virgílio Bento[1,3], PhD; Márcio Colunas[1], MSc; Nelson Rocha[2], PhD; Paula Coutinho[4], PhD

[1]Hospital São Sebastião, Neurology Department, Centro Hospitalar de Entre o Douro e Vouga, Santa Maria da Feira, Portugal

[2]Clinical Research Office, Health Sciences Department, University of Aveiro, Aveiro, Portugal

[3]University Institute of Maia, Maia, Portugal

[4]UnIGENe, Instituto de Biologia Molecular e Celular, University of Porto, Porto, Portugal

Corresponding Author:
Vítor Tedim Cruz, MD
Hospital São Sebastião
Neurology Department
Centro Hospitalar de Entre o Douro e Vouga
Rua Dr Candido de Pinho
Santa Maria da Feira, 4520-211
Portugal
Phone: 351 912582120
Fax: 351 256373867
Email: vitor.cruz@chedv.min-saude.pt

Abstract

Background: Cognitive training has been playing an increasing role in the treatment of patients with cognitive deficits. This type of intervention, namely its intensity, can be optimized by incorporating information technology-based systems.

Objective: The intent of the study was to determine the treatment intensity and patient adherence to home-based cognitive training strategies (Web-based cognitive training).

Methods: A cohort of 45 patients with neurologic and psychiatric diseases attending an outpatient memory clinic (average age 50.7 years, SD 17.0; average education 7.8 years, SD 4.9) was followed over 18 months. Participants were challenged to use a Web-based cognitive training system, “COGWEB”, on a daily basis, and fulfilled at least four weeks of training supervised remotely. Additionally, 11 patients attended face-to-face sessions.

Results: The average duration of continuous cognitive training was 18.8 weeks (SD 18.9). Each patient performed on average 363.5 minutes/week (SD 136.6). At 6-month follow-up, 82.8% complied with their treatment plan. The average proportion of complete weeks was 0.75 (SD 0.22). Patients with dementia trained more intensively (444.6 minutes/week), followed by patients with static brain lesion (414.5 minutes/week; *P*=.01). The group that held face-to-face sessions performed more training overall (481.4 vs 366.9 minutes/week), achieving a stronger expression and statistical significance in the last week of training (652.6 versus 354.9 minutes/week, *P*=.027).

Conclusions: Overall, the weekly training intensity was high. Patients with dementia and static lesions performed more cognitive training. Face-to-face sessions were associated with higher intensities. The combination of classical methods with information technology systems seems to ensure greater training intensity.

(J Med Internet Res 2014;16(5):e122) doi:10.2196/jmir.3377

KEYWORDS

cognitive training; neurorehabilitation; Web-based training; eHealth systems; training intensity; adherence; memory clinic

Introduction

Cognitive deficits are a common expression of highly prevalent neurological and psychiatric conditions that may affect individuals of all ages and usually have a long-lasting course [1]. This group of diseases includes Alzheimer’s and vascular dementias, stroke, Parkinson’s disease, traumatic brain injury, multiple sclerosis, bipolar disease, schizophrenia, attention

deficit hyperactivity disorder, and all sorts of developmental delays [1-4].

Health systems in general are developing more targeted approaches to these conditions, like adult memory clinics, developmental clinics, comprehensive rehabilitation centers, and community-based approaches, directed at either the older population with neurodegenerative diseases [5] or school-age children with learning disabilities [3,6]. All these strategies aim to improve care, mainly through a combination of prompt detection of cognitive deficits in populations at risk and early reference and therapeutic interventions. In spite of the huge efforts to organize and improve care, both for patients and their caregivers, most of these conditions share some ominous characteristics. They are chronic and to date have no substantial pharmacological treatments [7,8].

In this context, cognitive training has been playing an ever-increasing role in the treatment of patients with cognitive deficits. More and more studies have reported some beneficial effects of cognitive training in ageing [9], mild to moderate Alzheimer's disease and vascular dementia [10], Parkinson's disease [11], stroke and brain injury [12], multiple sclerosis [13,14], depression, or schizophrenia [15]. In addition, some data gathered also support the idea that improvements attributed to training may generalize beyond task-specific skills [16-18], but this remains controversial due to the lack of randomized trials with appropriate controls [10,19,20]. Mostly due to methodological issues, the evidence gathered is far from providing a clear demonstration of the benefits of cognitive training and much effort is warranted to improve the design of future interventions and trials [10,21-24].

In addition, scientific discussion in the field has been raising some additional questions: (1) how to deliver this type of treatment efficiently to larger numbers of patients in need of it, (2) how to monitor and control its effects over long periods of time in real-life clinical settings, and (3) how to accommodate the increasing knowledge of neuroplastic properties of the brain and future neuro-pharmacological tools [21,25,26].

Since the number of patients that could be eligible for this type of treatment is ever increasing, it is essential to develop and validate new strategies that may improve access without elevating the costs to deliver such care [6,27]. The incorporation of computers and information technology-based systems in our current practice may optimize cognitive interventions, namely their intensity, patient adherence, and quality of professional monitoring [28-31].

We have been working on a previously described Web-based cognitive training system, "COGWEB", since 2005. Over the years, its characteristics were tailored to address major needs identified in a memory clinic setting [32-34]. This clinic organizes and delivers care to a population of 400,000, and is based in a hospital institution with clinical and research activities.

With the present study, we aimed to analyze aspects of the quality of the cognitive training delivered, specifically, adherence and continued use of the training program in the most important subgroups of diseases attending an ordinary memory clinic setting. This was a follow-up study, focused on the investigation of the intensity of cognitive training achieved and patient adherence to treatment, using COGWEB to deliver home-based cognitive training over long periods of time.

Methods

Clinical Setting and Patient Selection

The study was based in a memory clinic that provides care to neurologic and psychiatric patients of all ages (adult and pediatric) with cognitive impairment, irrespective of their baseline disease. The resident staff members include neurologists and neuropsychologists, who collaborate with other departments in a tertiary hospital. Patients are referred to this clinic by other neurologists, neurosurgeons, psychiatrists, rehabilitation medicine physicians, pediatricians, internists, or general practitioners. From this outpatient memory clinic, consecutive patients that fulfilled all of the following inclusion criteria were selected: (1) medical diagnosis of a neurologic or psychiatric condition known to produce cognitive impairment, (2) cognitive deficits confirmed by comprehensive neuropsychological evaluation using tests validated for the Portuguese population, covering domains such as attention, memory, language, executive functions, and constructional ability and selected on the basis of pathology and patient characteristics (scores were reviewed by two senior neuropsychologists and each patient was classified as having or not having a deficit in each cognitive domain), (3) at least four years of formal education completed and ability to use personal computers and information technology applications, (4) favorable opinion of the attending physician and neuropsychologist toward enrollment in cognitive training activities, (5) no sensory or physical deficiency that could prevent the independent use of personal computers and information technology applications (eg, blindness, hemiplegia, or amputation), and (6) informed consent from both the patient and relative.

There were no limits of age for inclusion. Patients were first proposed by their attending physician for enrollment in cognitive rehabilitation strategies between July and December 2011. For data analysis, only the patients that had started their treatment at least four weeks before the end of the study (18 months after study beginning) were considered. This was done to guarantee a minimum follow-up time for the within-subjects adherence analysis. During the enrollment period, 240 patients were assessed at the clinic for the first time, of which 30 were classified as not having cognitive impairment. Of those remaining, 80 did not fulfill the required level of education or ability to use personal computers. Additionally, patients were deemed ineligible due to the severity of their disease or comorbidities (n=48), sensory or physical deficiency complicating stroke, diabetes, or cataracts (n=7), and no available relative to sign the informed consent (n=3).

Due to the heterogeneity of the conditions at this memory clinic [32], and to facilitate the analysis of data, patients were grouped according to their baseline pathology into four groups: (1) neurodegenerative diseases (eg, mild stages of Alzheimer's disease, frontotemporal dementia, or Parkinson's disease), (2)

memory complaints with depressive symptoms, (3) static brain lesions (eg, stroke, traumatic brain injury, or encephalitis), and (4) other diseases (eg, epilepsy, inflammatory diseases, schizophrenia, or attention deficit hyperactive disorder).

Ethical Issues

All patients and caregivers understood the purpose of the study and provided written informed consent. Approvals from the referring neurologists were also obtained to guarantee that the expectations of patients and caregivers were properly managed. This study was approved by the hospital review board and ethics commission (Hospital São Sebastião, Centro Hospitalar de Entre o Douro e Vouga, Santa Maria da Feira, Portugal).

Cognitive Intervention

Main Characteristics of COGWEB

The COGWEB system allows for the implementation of personalized cognitive training programs remotely, in the patient's living environment, under continuous supervision by experienced neuropsychologists [32]. The version used for this study was composed of 27 independent exercises in a computerized game format, developed to train various degrees of cognitive defects from mild to more severe impairments. Each exercise is organized primarily around a specific cognitive function, such as attention, executive functions, memory, language, praxis, gnosis, and calculus. Exercise progression is automatic through several levels of difficulty that change in accordance with the patient's performance and are coupled with support messages in real-time. The different degrees of difficulty are obtained through the manipulation of some features such as the number and type of items per level, their intrinsic complexity, or the interval between stimuli. All exercises use random, non-sequential stimuli to prevent memorization and maintain motivation between sessions. There are also several progress graphs (eg, right answers vs wrong answers, levels completed, global training time, or accesses) that are used to motivate patients after revision by the professional in charge [32,34].

Cognitive Training Design and Methods Used

The activities concerning cognitive training plans were all supervised by the resident neuropsychologist, who also conducted comprehensive neuropsychological assessments according to the patient medical diagnosis and using tests validated for the Portuguese population. All patients performed Web-based cognitive training, using the COGWEB system [32,34]. The training sessions were performed outside the hospital, predominantly at patients' homes or other comfortable family or social settings. The neuropsychologist tailored the cognitive training plan to the patients' medical conditions and cognitive deficits, thus contents of the training sessions varied during the course of the rehabilitation program. Sessions could include exposure to different combinations and proportions of exercises focused either on memory, executive functioning, attention, language, calculation, or constructive ability. The personalization of the cognitive training plans included the following possibilities (COGWEB system features): (1) recommended duration of each daily session, (2) number of sessions per week, (3) time of the day where most training should take place (morning or afternoon), (4) type, number, initial level of difficulty, and duration of each exercise (from a pool of 27) that composed the sessions, (5) frequency of adjustments to the exercises prescribed, and (6) frequency of progress reports from the neuropsychologist to the patient/caregiver. Patients were instructed to complete a minimum number of sessions per week (7 sessions, minimum of 30 minutes each). These could be performed at the patient and caregiver's convenience, at any time of the day in consecutive days or up to 4 sessions per day. Anything below this limit was considered non-adherence. There were no restrictions or indications of a maximum time of treatment per week.

Based on the clinical judgment of the neuropsychologists and attending physicians, some patients had their training programs based primarily on weekly face-to-face sessions with a neuropsychologist, either individualized or group sessions with an average duration of 60 minutes. Their internal organizations were defined by the neuropsychologists, according to each patient's baseline assessment and ongoing Web-based cognitive training activities. In the specific setting of the memory clinic where the study was based, face-to-face sessions are used primarily in the rehabilitation programs of younger patients with not only static brain lesions, which are usually more severe, but also with a higher potential for socioprofessional reintegration. Older patients with stroke and early dementia may also receive this type of treatment but mainly in group sessions.

Study Flow

In total, 72 patients fulfilled the inclusion criteria during the recruitment period. From these, 63 patients met all conditions that allowed them to start using the COGWEB system as part of their training program. Nonetheless, 8 patients (12.7%) did not actually start and 10 (15.6%) had used the system for a period of less than four weeks at the time of the analysis (Figure 1).

The analysis was conducted on a final sample of 45 patients with a mean age of 50.7 years (SD 17.0, range 11.0-84.0), mean years of formal education of 7.8 (SD 4.9, range 4.0-17.0), and 16 (35.6%) were female. According to their baseline pathology, of the 45 patients, 9 (20.0%) had definite neurodegenerative diseases, 14 (31.1%) had memory complaints with depressive symptoms, 15 (33.3%) had static brain lesions, and 7 (15.6%) had other diseases (Table 1). Patients that interrupted their treatment plan due to technical problems with the Internet at home or by their own decision were considered as non-adherent with treatment plan (Figure 1).

The 18 patients excluded from the analysis after agreeing to use COGWEB had a mean age of 49.0 (SD 17.4, range 19.0-78.0), mean years of formal education of 10.6 (SD 4.6, range 4.0-17.0), and 42% were female. Their baseline pathologies were: 22.2% (4/18) neurodegenerative diseases, 22.2% (4/18) memory complaints with depressive symptoms, 38.9% (7/18) static brain lesions, and 16.7% (3/18) other diseases.

Figure 1. Study flowchart.

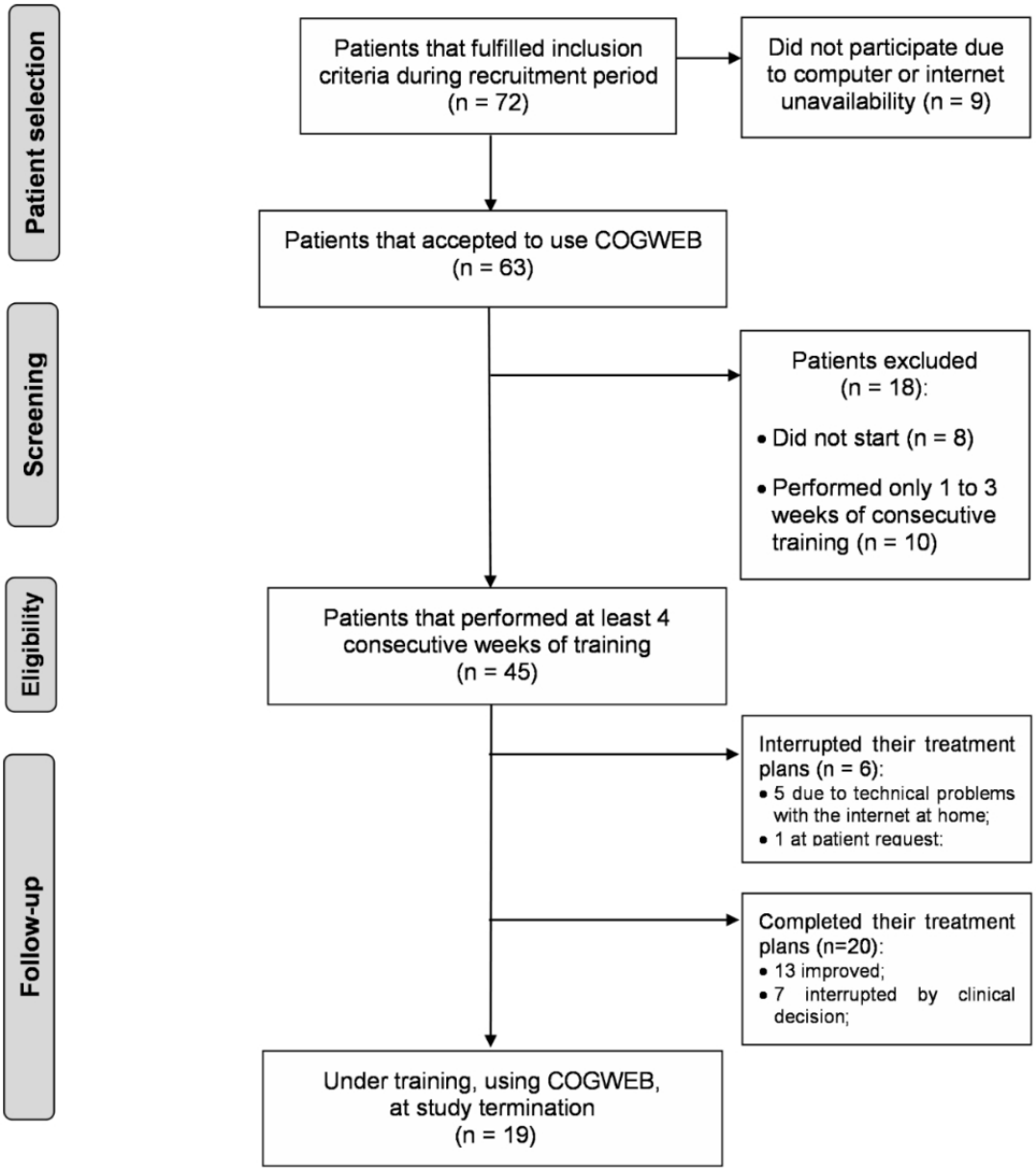


Table 1. Demographic characteristics of all groups.

Characteristics		Neuro-degenerative diseases, (n=9)	Memory complaints/ depression, (n=14)	Static brain lesions, (n=15)	Other diseases, (n=7)
Age (years), mean (SD)		61.8 (5.7)	54.8 (13.8)	44.2 (19.5)	44.6 (19.5)
Gender, n (%) male		8 (88.9)	6 (42.8)	11 (73.3)	4 (57.1)
Formal education (years), mean (SD)		6.5 (4.6)	6.1 (4.1)	9.2 (4.3)	9.9 (6.7)
Baseline cognitive performance, n (%) with deficit					
	Attention	9 (100.0)	12 (85.7)	13 (86.7)	4 (57.1)
	Memory	8 (88.9)	7 (50.0)	14 (93.3)	9 (100.0)
	Language	4 (44.4)	0 (0.0)	3 (20.0)	1 (14.3)
	Executive functioning	9 (100.0)	3 (21.4)	12 (80.0)	5 (71.4)
	Constructional ability	4 (44.4)	0 (0.0)	2 (13.3)	1 (14.3)
Face-to-face sessions, n (%) exposed		3 (33.3)	0 (0.0)	7 (46.7)	1 (14.3)

Outcome Definition

The COGWEB system allowed for the continuous monitoring of the following outcomes: (1) expected time of training (minutes)—summation of the duration of all prescribed sessions of training during the follow-up period of each patient, (2) time spent training (minutes)—summation of the duration of all sessions actually performed by the patient, (3) cumulative time of training in the first and last week of follow-up (minutes/week)—time of training in the first and last weeks, (4) assiduity—difference between the minimum number of sessions prescribed and the number of sessions actually performed, expressed as the proportion of complete weeks, and (5) follow-up period (weeks)—duration of consecutive time in training for each patient, with interruptions of more than one week duration being considered as study termination and the end of the follow-up period for a particular patient. This was further categorized as withdrawal due to non-adherence or termination according to treatment plan. The first two outcomes were used to measure the intensity of cognitive training obtained and the last three to measure motivation and adherence to treatment. Cognitive training plans were also classified as exclusively Web-based if all treatment activities occurred through the COGWEB, or combined when there was weekly face-to-face cognitive training work complemented with Web-based cognitive training activities.

Statistical Analysis

The SPSS Statistics version 21.0.0 software was used [35]. In order to characterize the global sample, mean values and standard deviations were used to describe outcomes, and parametric tests for statistical analysis were: ANOVA (analysis of variance), Student's *t* test for independent groups, and paired *t* test for within-subject comparison of cumulative time of training in the first and last week. For subgroup description, the median and interquartile ranges (IQR) were used as they are more suitable to the size and type of distribution within each group sample. To analyze the differences in outcomes between subgroups, the Kruskal-Wallis independent samples median test was used, adjusting for multiple comparisons. The related samples Wilcoxon signed-rank test was used to compare the first and the last weeks of training within each subgroup. The independent samples Mann-Whitney *U* test was used to compare the main demographic characteristics and the outcome differences between the group with exclusive Web-based training and the group with face-to-face sessions complemented with Web-based training. Fisher's exact test and chi-square were used to compare baseline characteristics such as gender, distribution of groups of diseases, and cognitive domains impaired, between subgroups. The effect of face-to-face sessions within subgroups of diseases was not analyzed due to the reduced sample size. Finally, an analysis of the probability to comply with the Web-based cognitive training was conducted using the Kaplan-Meier survival method in order to model the duration time of the treatment up to its interruption. Patients completing the treatment plan or undergoing training at the time of the follow-up were censored.

Results

Intensity of Treatment Obtained

For the duration of the entire follow-up period, patients performed on average 363.5 minutes/week (SD 136.6, range 84.7-652.6) of cognitive training activities through the COGWEB system. This was 1.7 times higher than the minimum requirement.

The analysis of the mean time training per week between groups of diseases revealed significant differences (Figure 2 and Table 2), with neurodegenerative diseases and static brain lesions dedicating more time to training (H_3=11.41, P=.01). There was no association of mean time training per week with potential confounders like age ($F_{1,41}$=0.86, P=.36), gender (t_{42}=−1.64, P=.11) or education ($F_{1,41}$= 0.70, P=.41).

Figure 2. Time spent training (average in minutes/week) per disease group.

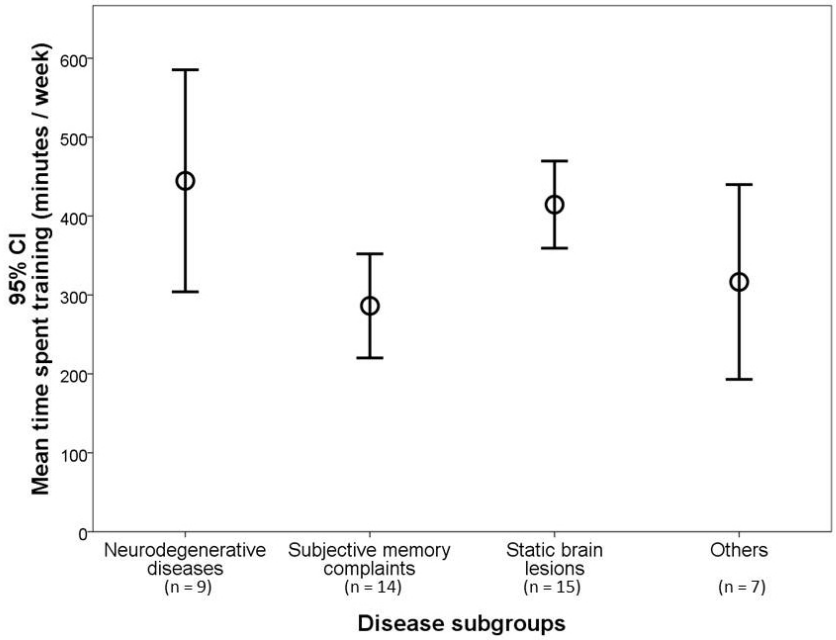


Table 2. Indicators of intensity and adherence to treatment per major group of diseases.

	Neurodegenerative diseases, median (interquartile range)	Memory complaints/ depression, median (interquartile range)	Static brain lesions, median (interquartile range)	Other diseases, median (interquartile range)
Follow-up duration (weeks)	26 (7.8-29.8)	22.5 (7.8-33.5)	8 (5.0-25.0)	11 (7.0-18.0)
Time training per week (minutes)	479.0 (257.6-567.7)	295.3 (187.3-404.0)	423.6 (362-458)	295.6 (203.7-366.9)
Time training, first week (minutes)	555.9 (159.0-806.0)	308.3 (143.2-579.8)	501.9 (442.9-656.3)	173.3 (99.7-491.8)
Time training, last week (minutes)	394.6 (201.0-639.4)	282.5 (73.3-576.2)	376.0 (279.8-804.8)	379.5 (254.3-443.2)
Assiduity (proportion of complete weeks)	0.89 (0.53-0.96)	0.73 (0.55-0.84)	0.80 (0.75-1.0)	0.63 (0.53-0.83)

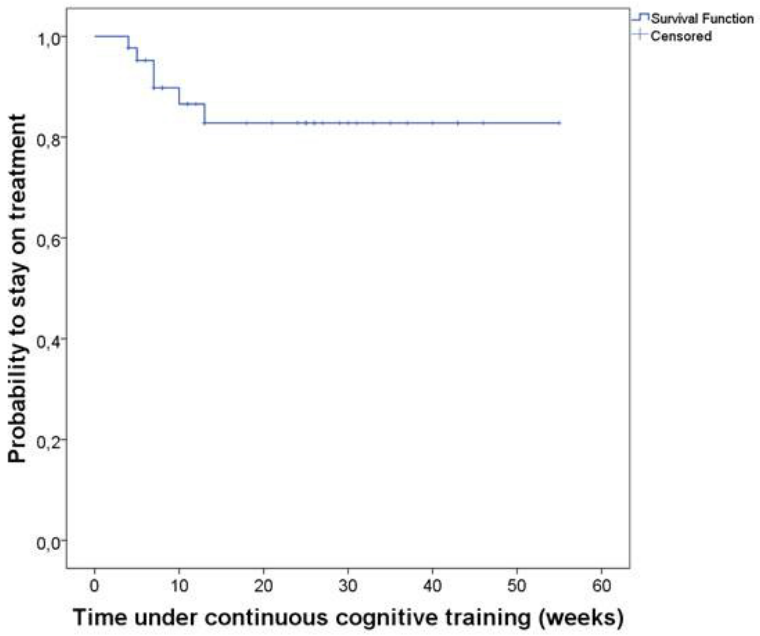
Adherence to Treatment

The average duration of continuous cognitive training was 18.8 weeks (SD 18.9, range 4.0-55.0), and there were no statistically significant differences among groups (H_3=3.40, P=.33) (Table 2). During the first week, the average time training was 428.7 minutes (SD 264.8, range 21.0-891.0). In the final week, this value was 414.5 minutes (SD 268.1, range 21.1-969.0). These values were not statistically different (t_{43}=0.27, P=.79). There were no differences of mean time training between first and last week attributable to any of the major group of diseases (Z=22.00, P=.58 for neurodegenerative diseases; Z=53.00, P=.98 for memory complaints with depression; Z=63.00, P=.87 for static brain lesions; Z=14.00, P=1.00 for other diseases) (Table 2).

The average proportion of complete weeks of training (measure of assiduity) was 0.75 (SD 0.22, range 0.18-1.0) and there were no difference between groups (H_3=4.04, P=.26) (Table 2).

The application of the Kaplan-Meier method estimated an average duration of continuous Web-based cognitive treatment of 46.9 weeks (SD 3.03), with 95% confidence intervals of 41.3 and 52.8 weeks. At 6-month follow-up (24 weeks), 82.8% of patients complied with their treatment plan (Figure 3).

Figure 3. Probability of continuing with treatment over time (Kaplan-Meier survival function) for the first 60 weeks. There were no treatment interruptions after this period. Patients completing the treatment plan or undergoing training at time of follow-up were censored.



Impact of Face-to-Face Sessions

During the follow-up period, 11/45 patients (24.4%) received weekly face-to-face sessions complemented with Web-based training (63.6%, 7/11 static brain lesions, 27.3%, 3/11 neurodegenerative, and 9.1%, 1/11 other diseases). Patients with memory complaints and depressive symptoms were excluded from this analysis since none in this subgroup was exposed to face-to-face sessions (Table 1). The baseline characteristics of the two groups are depicted in Table 3. There were no significant differences regarding age (U_{28}=123.0, P=.425), formal education (U_{28}=286.5, P=.718), gender (χ^2_1=0.6, P=.42), and distribution of the groups of diseases

(χ^2_2=1.8, *P*=.42) between the two groups. The distribution of cognitive impairment by domain was also similar (Table 3).

Table 3. Demographic characteristics of the groups used for analysis of the impact of face-to-face sessions.

Characteristics		Exclusively Web-based training (n=20)	Face-to-face sessions complemented with Web-based training (n=11)
Age (years), mean (SD)		50.0 (19.9)	47.2 (15.6)
Gender, n (%) male		13 (65.0)	9 (81.8)
Formal education (years), mean (SD)		8.5 (5.2)	8.9 (5.1)
Major groups of diseases, n (%)			
	ND[a]	6 (30.0)	3 (27.3)
	SBL[b]	8 (40.0)	7 (63.6)
	OD[c]	6 (30.0)	1 (9.1)
Baseline cognitive performance, n (%) with deficit			
	Attention	17 (85.0)	9 (81.8)
	Memory	18 (90.0)	11 (100.0)
	Language	5 (25.0)	3 (27.3)
	Executive functioning	16 (80.0)	10 (90.9)
	Constructional ability	4 (20.0)	3 (27.3)

[a]ND: neurodegenerative diseases

[b]SBL: static brain lesions

[c]OD: other diseases

The median duration of the follow-up was higher in the group with face-to-face sessions: 26.0 weeks (IQR=7.0−43.0; min. 4.0, max. 55.0) vs 11.0 weeks (IQR=6.0−18.0; min. 4.0, max. 40.0) in the group with exclusively Web-based training. However, there was no statistical significance (U_{28}=70.5, *P*=.145) (Table 4). The overall median time training per week in the group with face-to-face sessions was 481.4 minutes (IQR=398.4−577.3; min. 180.4, max. 652.6), while in the group with exclusively Web-based sessions it was 366.9 minutes (IQR=281.3−452.5; min. 191.3, max. 583.0). This difference had no statistical significance (U_{28}=62.0, *P*=.07). In the last week of the cognitive intervention, significant differences were verified in the median time training between the two groups with 652.6 minutes (IQR=379.5−817.4; min. 279.8, max. 969.0) when there were face-to-face sessions vs 354.9 minutes (IQR=138.5–577.3; min. 21.1, max. 857.0) when exclusively Web-based (U_{28}= 53.0, *P*=.027). These differences were not present in the first week of training (U_{28}=106.0, *P*=.949) (Table 4). The overall assiduity was not different between these two groups during the study (U_{28}=82.0, *P*=.33).

Table 4. Indicators of intensity and adherence to treatment per major type of treatment strategy.

	Exclusively Web-based training (n=20), median (interquartile range)	Face-to-face sessions complemented with Web-based training (n=11), median (interquartile range)
Follow-up duration (weeks)	11.0 (6.0-18.0)	26.0 (7.0-43.0)
Time training per week (minutes)	366.9 (281.3-452.5)	481.4 (398.4-577.3)
Time training, first week (minutes)	489.3 (145.9-662.9)	490.7 (173.3-655.2)
Time training, last week (minutes)	354.9 (138.5-577.3)	652.6 (379.5-817.4)
Assiduity (proportion of complete weeks)	0.75 (0.3-1.0)	0.83 (0.4-1.0)

Discussion

Principal Findings

This study provided data on the characteristics of cognitive training treatments using a Web-based approach in an ordinary memory clinic setting. The overall intensities of training obtained were very high, averaging 6 hours per week and exceeding 1.7 times of what was set as minimum. Furthermore, the characteristics of the system used (COGWEB) permitted uninterrupted training activities over long periods of time, with 82.8% of patients complying with treatment at 6 months. The combination of high intensity and long duration of treatment is very important to stimulate neuroplasticity in the brain [21], more so, if we consider the design of future randomized clinical

trials to assess the impact of cognitive training on functional outcomes in several important diseases [21,23,36].

Significant differences were found in the mean intensity of treatment obtained between groups, with neurodegenerative diseases and static brain injury performing around 7 hours of training per week, while people with memory complaints and depressive symptoms trained close to 5 hours per week. It is important to point out that all groups performed above the minimum requirements of 30 minutes of training per day (same for all). Engaging psychiatric or neurologic patients in training or interesting leisure activities is very difficult [37]. As an example of the current state of the art, even in inpatient mental health services of developed countries, the level of activities, other than sleep, eating, or watching TV, is less than 17 minutes per day [37]. This is in high contrast with what was obtained in this study for the several groups of diseases analyzed.

During the follow-up period of the 45 patients included, and specifically comparing the first and the last week of training, the intensity of treatment did not decay and there were no important effects attributable to the major disease groups. Furthermore, follow-up duration between major groups of diseases did not differ. Although neurodegenerative disease patients had a tendency to have longer follow-up periods (around 7 months), this could be explained only by clinical reasons, with static brain lesions being prescribed shorter periods of training. These latter findings may be due to the reduced sample size for subgroup analysis.

An interesting finding of this study was the effect of weekly face-to-face sessions on the overall intensities of Web-based cognitive training activities. The group exposed to face-to-face sessions performed, on average, 2 additional hours of training per week during the entire duration of the follow-up period. This difference was not present in the first week of training, but was built over time and achieved a value of 4 hours and statistical significance in the last week of training. There was a trend for longer follow-up periods in the group with face-to-face sessions, but not achieving statistical significance. These findings are in accordance with some critical analysis of the impact of computerized cognitive training activities and the need to prevent excessive isolation of patients during treatment [38-42]. In future studies, if the intensity of treatment and adherence are to be maximized, the inclusion of some kind of periodic face-to-face individual or group session is warranted. Nonetheless, to clarify the impact of different methods of face-to-face sessions (eg, individual, group, weekly, monthly) and whether they are reproducible between groups of diseases, further studies are necessary.

Limitations

The limitations of this study are mainly inherent to the uncontrolled nature and single center design, which impose some restrictions on the generalizability of the findings. In this respect, it is important to note that from the 240 patients initially assessed, 80 (33.3%) did not fulfil the required levels of literacy or ability to use personal computers and information technology applications. Furthermore, among the patients that fulfilled inclusion criteria, 9 out of 72 did not participate due to personal computer or Internet unavailability and 8 out of 63 did not start after agreeing to participate. These values may reflect the low literacy levels and barriers in patient access to information technology at home, in this segment of the Portuguese population [43]. Although the trends are changing [44], these aspects are still significant in the population aged over 50 and must be taken into consideration in the implementation of this type of cognitive intervention in clinical practice or future research.

In addition, the focus of this work was on obtaining data on the intensity and adherence to treatment and for that reason blinded information on cognitive baseline or outcome measures was not collected. The patient's diagnosis only conveys indirect information on patient deficits and level of impairment, with baseline cognitive performance data provided only partially addressing this limitation. Despite the inclusion criteria defined, the enrollment of patients in the study was based upon a referral by their attending physician and neuropsychologist's judgment. They decided whether the patient would comply with treatment and also if the deficits and background literacy or cognitive reserve were suitable. Face-to-face sessions were also decided on clinical indication and not randomized. The role of the professionals in patient selection in both these situations may have biased the results in a direction consistent with the findings. Furthermore, differences between first and last week intensities may also be due to selection biases attributable to the professional intervention. The heterogeneity of diagnoses was also a potential weakness and should not be maintained in trials evaluating clinical efficacy.

Future studies must analyze the impact of up to 7 hours of cognitive training per week on global motor activities, sedentarism indexes [45], and also possible negative mental effects of uncontrolled cognitive training activities [46]. These latter aspects are similar to the risks associated with unsupervised "of-the-shelf" home rehabilitation activities and learned non-use models during aphasia or motor rehabilitation after stroke [47-49]. They may only be avoided through control of several aspects of training like activities preformed, cumulative dose of training in each cognitive domain, and specific cognitive outcomes along time.

Conclusions

Overall, the training intensity achieved per week was high. The groups of patients with dementia and static lesions performed more cognitive training. Patients with additional face-to-face sessions achieved a higher intensity workout. The combination of classical methods with information technology-based systems like COGWEB seems to be the option that ensures greater training intensity. This method should be further explored in multicenter randomized controlled trials targeted at the most prevalent diseases like dementia, stroke, schizophrenia, or multiple sclerosis.

Acknowledgments

Several psychology students and administrative personnel at the hospital collaborated in the study, for which we are very grateful. We also thank our patients, their families, and their attending physicians for their collaboration.

The study was funded in part by research grants from the Portuguese Society of Neurology in 2009 and Grünenthal Foundation in 2010. This financial support was used for the technological development of COGWEB. All other design and conduct of the study, collection, management, analysis and interpretation of the data and preparation, review or approval of the manuscript received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Authors' Contributions

Cruz, Pais, and Coutinho obtained funding and were responsible for study concept and design. Cruz, Pais, Alves, Mateus, Ruano, Barreto, Bento, and Colunas were responsible for acquisition of data. Analysis and interpretation of data was performed by Cruz, Pais, Ruano, Rocha, and Coutinho. All authors participated in critical revision of the manuscript for important intellectual content. Administrative, technical, and material support was provided by Alves, Ruano, Mateus, Barreto, Bento, and Colunas, and study supervision was performed by Cruz, Pais, Bento, Colunas, Rocha, and Coutinho.

Conflicts of Interest

Cruz and Pais have a shareholder position at Neuroinova, Lda, a company that develops and commercializes COGWEB-related products. Bento and Colunas received fees for the technological development of COGWEB. Alves, Ruano, Mateus, Barreto, Rocha, and Coutinho have no conflicts of interest to report.

References

1. World Health Organization. Neurological disorders. Public health challenges. Geneva: World Health Organization; 2006.
2. Wimo A, Jönsson L, Bond J, Prince M, Winblad B, Alzheimer Disease International. The worldwide economic impact of dementia 2010. Alzheimers Dement 2013 Jan;9(1):1-11.e3. [doi: 10.1016/j.jalz.2012.11.006] [Medline: 23305821]
3. Flisher AJ, Hatherill S, Dhansay Y. Specific mental health disorders: Child and adolescent mental disorders. In: Kris H, editor. International encyclopedia of public health. Oxford: Academic Press; 2008:147-154.
4. Patel V, Flisher AJ, Hetrick S, McGorry P. Mental health of young people: a global public-health challenge. Lancet 2007 Apr 14;369(9569):1302-1313. [doi: 10.1016/S0140-6736(07)60368-7] [Medline: 17434406]
5. Ciolli L, Poggesi A, Salvadori E, Valenti R, Nannucci S, Pasi M, et al. The VAS-COG clinic: an out-patient service for patients with cognitive and behavioral consequences of cerebrovascular diseases. Neurol Sci 2012 Dec;33(6):1277-1283. [doi: 10.1007/s10072-012-0941-0] [Medline: 22258361]
6. Eaton J, McCay L, Semrau M, Chatterjee S, Baingana F, Araya R, et al. Scale up of services for mental health in low-income and middle-income countries. Lancet 2011 Oct 29;378(9802):1592-1603. [doi: 10.1016/S0140-6736(11)60891-X] [Medline: 22008429]
7. Anand R, Gill KD, Mahdi AA. Therapeutics of Alzheimer's disease: Past, present and future. Neuropharmacology 2014 Jan;76 Pt A:27-50. [doi: 10.1016/j.neuropharm.2013.07.004] [Medline: 23891641]
8. Rafii MS. Update on Alzheimer's disease therapeutics. Rev Recent Clin Trials 2013 Jun;8(2):110-118. [Medline: 23859061]
9. Rebok GW, Ball K, Guey LT, Jones RN, Kim HY, King JW, et al. Ten-year effects of the advanced cognitive training for independent and vital elderly Cognitive Training Trial on Cognition and Everyday Functioning in Older Adults. J Am Geriatr Soc 2014 Jan 13:16-24. [doi: 10.1111/jgs.12607] [Medline: 24417410]
10. Bahar-Fuchs A, Clare L, Woods B. Cognitive training and cognitive rehabilitation for persons with mild to moderate dementia of the Alzheimer's or vascular type: a review. Alzheimers Res Ther 2013 Aug 7;5(4):35 [FREE Full text] [doi: 10.1186/alzrt189] [Medline: 23924584]
11. Edwards JD, Hauser RA, O'Connor ML, Valdés EG, Zesiewicz TA, Uc EY. Randomized trial of cognitive speed of processing training in Parkinson disease. Neurology 2013 Oct 8;81(15):1284-1290. [doi: 10.1212/WNL.0b013e3182a823ba] [Medline: 24014503]
12. Chung CS, Pollock A, Campbell T, Durward BR, Hagen S. Cognitive rehabilitation for executive dysfunction in adults with stroke or other adult non-progressive acquired brain damage. Cochrane Database Syst Rev 2013;4:CD008391. [doi: 10.1002/14651858.CD008391.pub2] [Medline: 23633354]
13. Rosti-Otajärvi EM, Hämäläinen PI. Neuropsychological rehabilitation for multiple sclerosis. Cochrane Database Syst Rev 2011(11):CD009131. [doi: 10.1002/14651858.CD009131.pub2] [Medline: 22071863]
14. das Nair R, Ferguson H, Stark DL, Lincoln NB. Memory rehabilitation for people with multiple sclerosis. Cochrane Database Syst Rev 2012;3:CD008754. [doi: 10.1002/14651858.CD008754.pub2] [Medline: 22419337]
15. Vinogradov S, Fisher M, de Villers-Sidani E. Cognitive training for impaired neural systems in neuropsychiatric illness. Neuropsychopharmacology 2012 Jan;37(1):43-76 [FREE Full text] [doi: 10.1038/npp.2011.251] [Medline: 22048465]
16. Mackey AP, Hill SS, Stone SI, Bunge SA. Differential effects of reasoning and speed training in children. Dev Sci 2011 May;14(3):582-590. [doi: 10.1111/j.1467-7687.2010.01005.x] [Medline: 21477196]

17. Jaeggi SM, Studer-Luethi B, Buschkuehl M, Su Y, Jonides J, Perrig WJ. The relationship between n-back performance and matrix reasoning — implications for training and transfer. Intelligence 2010 Nov 2;38(6):625-635. [doi: 10.1016/j.intell.2010.09.001] [Medline: 23731643]
18. Jaeggi SM, Buschkuehl M, Jonides J, Perrig WJ. Improving fluid intelligence with training on working memory. Proc Natl Acad Sci U S A 2008 May 13;105(19):6829-6833 [FREE Full text] [doi: 10.1073/pnas.0801268105] [Medline: 18443283]
19. Diamond A, Lee K. Interventions shown to aid executive function development in children 4 to 12 years old. Science 2011 Aug 19;333(6045):959-964 [FREE Full text] [doi: 10.1126/science.1204529] [Medline: 21852486]
20. Papp KV, Walsh SJ, Snyder PJ. Immediate and delayed effects of cognitive interventions in healthy elderly: a review of current literature and future directions. Alzheimers Dement 2009 Jan;5(1):50-60. [doi: 10.1016/j.jalz.2008.10.008] [Medline: 19118809]
21. Cramer SC, Sur M, Dobkin BH, O'Brien C, Sanger TD, Trojanowski JQ, et al. Harnessing neuroplasticity for clinical applications. Brain 2011 Jun;134(Pt 6):1591-1609 [FREE Full text] [doi: 10.1093/brain/awr039] [Medline: 21482550]
22. Rabipour S, Raz A. Training the brain: fact and fad in cognitive and behavioral remediation. Brain Cogn 2012 Jul;79(2):159-179. [doi: 10.1016/j.bandc.2012.02.006] [Medline: 22463872]
23. Langenbahn DM, Ashman T, Cantor J, Trott C. An evidence-based review of cognitive rehabilitation in medical conditions affecting cognitive function. Arch Phys Med Rehabil 2013 Feb;94(2):271-286. [doi: 10.1016/j.apmr.2012.09.011] [Medline: 23022261]
24. Sonuga-Barke EJ, Brandeis D, Cortese S, Daley D, Ferrin M, Holtmann M, European ADHD Guidelines Group. Nonpharmacological interventions for ADHD: systematic review and meta-analyses of randomized controlled trials of dietary and psychological treatments. Am J Psychiatry 2013 Mar 1;170(3):275-289. [doi: 10.1176/appi.ajp.2012.12070991] [Medline: 23360949]
25. McLaren AN, Lamantia MA, Callahan CM. Systematic review of non-pharmacologic interventions to delay functional decline in community-dwelling patients with dementia. Aging Ment Health 2013 Aug;17(6):655-666. [doi: 10.1080/13607863.2013.781121] [Medline: 23611141]
26. Aguirre E, Woods RT, Spector A, Orrell M. Cognitive stimulation for dementia: a systematic review of the evidence of effectiveness from randomised controlled trials. Ageing Res Rev 2013 Jan;12(1):253-262. [doi: 10.1016/j.arr.2012.07.001] [Medline: 22889599]
27. Gates NJ, Sachdev PS, Fiatarone Singh MA, Valenzuela M. Cognitive and memory training in adults at risk of dementia: a systematic review. BMC Geriatr 2011;11:55 [FREE Full text] [doi: 10.1186/1471-2318-11-55] [Medline: 21942932]
28. Cha YJ, Kim H. Effect of computer-based cognitive rehabilitation (CBCR) for people with stroke: a systematic review and meta-analysis. NeuroRehabilitation 2013;32(2):359-368. [doi: 10.3233/NRE-130856] [Medline: 23535800]
29. Minzenberg MJ, Carter CS. Developing treatments for impaired cognition in schizophrenia. Trends Cogn Sci 2012 Jan;16(1):35-42. [doi: 10.1016/j.tics.2011.11.017] [Medline: 22178120]
30. Kueider AM, Parisi JM, Gross AL, Rebok GW. Computerized cognitive training with older adults: a systematic review. PLoS One 2012;7(7):e40588 [FREE Full text] [doi: 10.1371/journal.pone.0040588] [Medline: 22792378]
31. Simon SS, Yokomizo JE, Bottino CM. Cognitive intervention in amnestic Mild Cognitive Impairment: a systematic review. Neurosci Biobehav Rev 2012 Apr;36(4):1163-1178. [doi: 10.1016/j.neubiorev.2012.01.007] [Medline: 22322184]
32. Cruz VT, Pais J, Bento V, Mateus C, Colunas M, Alves I, et al. A rehabilitation tool designed for intensive Web-based cognitive training: description and usability study. JMIR Res Protoc 2013;2(2):e59 [FREE Full text] [doi: 10.2196/resprot.2899] [Medline: 24334248]
33. Cruz VT, Pais J, Bento VF, Mateus C, Colunas M, Alves I. Neuroinova, Lda. 2013. Cogweb® Web-based cognitive training platform URL: http://www.cogweb.eu/ [accessed 2013-08-22] [WebCite Cache ID 6J3yBxQrd]
34. Cruz VT, Pais J. Cogweb® – sistema integrado de estimulação cognitiva: Manual de formação para profissionais. Gaia: Neuroinova; 2012.
35. IBM Corp. IBM SPSS Statistics for Windows. Armonk, NY: IBM Corp; 2012.
36. Langhorne P, Bernhardt J, Kwakkel G. Stroke rehabilitation. Lancet 2011 May 14;377(9778):1693-1702. [doi: 10.1016/S0140-6736(11)60325-5] [Medline: 21571152]
37. Killaspy H, Cook S, Mundy T, Craig T, Holloway F, Leavey G, et al. Study protocol: cluster randomised controlled trial to assess the clinical and cost effectiveness of a staff training intervention in inpatient mental health rehabilitation units in increasing service users' engagement in activities. BMC Psychiatry 2013 Aug 28;13(1):216 [FREE Full text] [doi: 10.1186/1471-244X-13-216] [Medline: 23981710]
38. George DR, Whitehouse PJ. Marketplace of memory: what the brain fitness technology industry says about us and how we can do better. Gerontologist 2011 Oct;51(5):590-596. [doi: 10.1093/geront/gnr042] [Medline: 21572161]
39. Bavelier D, Davidson RJ. Brain training: Games to do you good. Nature 2013 Feb 28;494(7438):425-426. [doi: 10.1038/494425a] [Medline: 23446401]
40. Bavelier D, Green CS, Han DH, Renshaw PF, Merzenich MM, Gentile DA. Brains on video games. Nat Rev Neurosci 2011 Dec;12(12):763-768. [doi: 10.1038/nrn3135] [Medline: 22095065]

41. Hedges JH, Adolph KE, Amso D, Bavelier D, Fiez JA, Krubitzer L, et al. Play, attention, and learning: how do play and timing shape the development of attention and influence classroom learning? Ann N Y Acad Sci 2013 Jul;1292:1-20 [FREE Full text] [doi: 10.1111/nyas.12154] [Medline: 23763338]
42. Cicerone KD, Langenbahn DM, Braden C, Malec JF, Kalmar K, Fraas M, et al. Evidence-based cognitive rehabilitation: updated review of the literature from 2003 through 2008. Arch Phys Med Rehabil 2011 Apr;92(4):519-530. [doi: 10.1016/j.apmr.2010.11.015] [Medline: 21440699]
43. Nunes B, Silva RD, Cruz VT, Roriz JM, Pais J, Silva MC. Prevalence and pattern of cognitive impairment in rural and urban populations from Northern Portugal. BMC Neurol 2010;10:42 [FREE Full text] [doi: 10.1186/1471-2377-10-42] [Medline: 20540726]
44. Kummervold PE, Chronaki CE, Lausen B, Prokosch HU, Rasmussen J, Santana S, et al. eHealth trends in Europe 2005-2007: a population-based survey. J Med Internet Res 2008;10(4):e42 [FREE Full text] [doi: 10.2196/jmir.1023] [Medline: 19017584]
45. Rhodes RE, Mark RS, Temmel CP. Adult sedentary behavior: a systematic review. Am J Prev Med 2012 Mar;42(3):e3-28. [doi: 10.1016/j.amepre.2011.10.020] [Medline: 22341176]
46. Iuculano T, Cohen Kadosh R. The mental cost of cognitive enhancement. J Neurosci 2013 Mar 6;33(10):4482-4486 [FREE Full text] [doi: 10.1523/JNEUROSCI.4927-12.2013] [Medline: 23467363]
47. Sickert A, Anders LC, Münte TF, Sailer M. Constraint-induced aphasia therapy following sub-acute stroke: a single-blind, randomised clinical trial of a modified therapy schedule. J Neurol Neurosurg Psychiatry 2014 Jan;85(1):51-55. [doi: 10.1136/jnnp-2012-304297] [Medline: 23828834]
48. Taub E. The behavior-analytic origins of constraint-induced movement therapy: an example of behavioral neurorehabilitation. Behav Anal 2012;35(2):155-178 [FREE Full text] [Medline: 23449867]
49. Meinzer M, Elbert T, Djundja D, Taub E, Rockstroh B. Extending the Constraint-Induced Movement Therapy (CIMT) approach to cognitive functions: Constraint-Induced Aphasia Therapy (CIAT) of chronic aphasia. NeuroRehabilitation 2007;22(4):311-318. [Medline: 17971622]

Abbreviations

ANOVA: analysis of variance
COGWEB: cognitive Web-based training system
IQR: interquartile range

Edited by G Eysenbach; submitted 04.03.14; peer-reviewed by F Schmiedek, K Cicerone, K Ganci; comments to author 30.03.14; revised version received 22.04.14; accepted 28.04.14; published 07.05.14.

Please cite as:
Tedim Cruz V, Pais J, Alves I, Ruano L, Mateus C, Barreto R, Bento V, Colunas M, Rocha N, Coutinho P
Web-Based Cognitive Training: Patient Adherence and Intensity of Treatment in an Outpatient Memory Clinic
J Med Internet Res 2014;16(5):e122
URL: http://www.jmir.org/2014/5/e122/
doi:10.2196/jmir.3377
PMID:24808451

Original Paper

Monitoring of Internet Forums to Evaluate Reactions to the Introduction of Reformulated OxyContin to Deter Abuse

Emily C McNaughton[1], MPH; Paul M Coplan[2,3], ScD; Ryan A Black[4], PhD; Sarah E Weber[1], BS; Howard D Chilcoat[2,5], ScD; Stephen F Butler[1], PhD

[1]Inflexxion, Inc., Newton, MA, United States

[2]Purdue Pharma L.P., Stamford, CT, United States

[3]Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

[4]Center for Psychological Studies, Nova Southeastern University, Fort Lauderdale, FL, United States

[5]Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

Corresponding Author:
Emily C McNaughton, MPH
Inflexxion, Inc.
320 Needham Street, Suite 100
Newton, MA, 02464
United States
Phone: 1 617 614 6028 ext 256
Fax: 1 617 332 1820
Email: emcnaughton@inflexxion.com

Abstract

Background: Reformulating opioid analgesics to deter abuse is one approach toward improving their benefit-risk balance. To assess sentiment and attempts to defeat these products among difficult-to-reach populations of prescription drug abusers, evaluation of posts on Internet forums regarding reformulated products may be useful. A reformulated version of OxyContin (extended-release oxycodone) with physicochemical properties to deter abuse presented an opportunity to evaluate posts about the reformulation in online discussions.

Objective: The objective of this study was to use messages on Internet forums to evaluate reactions to the introduction of reformulated OxyContin and to identify methods aimed to defeat the abuse-deterrent properties of the product.

Methods: Posts collected from 7 forums between January 1, 2008 and September 30, 2013 were evaluated before and after the introduction of reformulated OxyContin on August 9, 2010. A quantitative evaluation of discussion levels across the study period and a qualitative coding of post content for OxyContin and 2 comparators for the 26 month period before and after OxyContin reformulation were conducted. Product endorsement was estimated for each product before and after reformulation as the ratio of endorsing-to-discouraging posts (ERo). Post-to-preintroduction period changes in EROs (ie, ratio of EROs) for each product were also calculated. Additionally, post content related to recipes for defeating reformulated OxyContin were evaluated from August 9, 2010 through September 2013.

Results: Over the study period, 45,936 posts related to OxyContin, 18,685 to Vicodin (hydrocodone), and 23,863 to Dilaudid (hydromorphone) were identified. The proportion of OxyContin-related posts fluctuated between 6.35 and 8.25 posts per 1000 posts before the reformulation, increased to 10.76 in Q3 2010 when reformulated OxyContin was introduced, and decreased from 9.14 in Q4 2010 to 3.46 in Q3 2013 in the period following the reformulation. The sentiment profile for OxyContin changed following reformulation; the post-to-preintroduction change in the ERo indicated reformulated OxyContin was discouraged significantly more than the original formulation (ratio of EROs=0.43, $P<.001$). A total of 37 recipes for circumventing the abuse-deterrent characteristics of reformulated OxyContin were observed; 32 were deemed feasible (ie, able to abuse). The frequency of posts reporting abuse of reformulated OxyContin via these recipes was low and decreased over time. Among the 5677 posts mentioning reformulated OxyContin, 825 posts discussed recipes and 498 reported abuse of reformulated OxyContin by such recipes (41 reported injecting and 128 reported snorting).

Conclusions: After introduction of physicochemical properties to deter abuse, changes in discussion of OxyContin on forums occurred reflected by a reduction in discussion levels and endorsing content. Despite discussion of recipes, there is a relatively

small proportion of reported abuse of reformulated OxyContin via recipes, particularly by injecting or snorting routes. Analysis of Internet discussion is a valuable tool for monitoring the impact of abuse-deterrent formulations.

(J Med Internet Res 2014;16(5):e119) doi:10.2196/jmir.3397

KEYWORDS

Internet; opioid analgesic; drug abuse; prescription drug; OxyContin; epidemiology; surveillance; social media; qualitative research

Introduction

Prescription opioid analgesics are an important component of pain management. Misuse and abuse of these medications, however, have created a serious and growing public health problem [1]. The balance between providing access to and prescribing these medications for patients with chronic pain while minimizing their diversion and abuse remains a significant challenge for all stakeholders, including prescribers, pharmaceutical manufacturers, and the Food and Drug Administration [2,3]. One important step toward the goal of creating safer opioid analgesics has been the development of opioid formulations designed to deter abuse [4-6]. These formulations are commonly referred to as abuse-deterrent formulations (ADFs) [7] or tamper-resistant formulations (TRFs). The science of deterring abuse via these formulations is new, and both the formulation technologies and the analytical, clinical, epidemiological, and statistical methodology for evaluating those technologies are rapidly evolving.

Most abuse-deterrent technologies developed to date are designed to make product manipulation more difficult or to make abuse of the manipulated product less attractive or rewarding. Although in vitro and clinical studies indicate the efficacy of these technologies, postmarketing data are needed to evaluate their effectiveness. One of the early formulations intended to reduce abuse was a reformulated version of extended-release oxycodone (reformulated OxyContin, Purdue Pharma, Stamford, CT, USA), which was introduced to the market in August 2010. This product has physicochemical resistance to crushing and dissolution intended to present obstacles to abuse by nonoral routes of administration (ROA) (eg, injecting, snorting). The launch of reformulated OxyContin provided a nationwide experiment to evaluate the impact of a product intended to reduce tampering in the real world [8,9]. To date, evidence from individuals evaluated for treatment triage suggests that reformulated OxyContin results in lower rates of abuse through nonoral abuse and abuse via any ROA [8] compared to historical rates for the original formulation of OxyContin. These findings, as well as others [10,11] that suggest reformulated OxyContin inhibits manipulation and abuse, are based on reports by abusers to some authority (eg, researcher, treatment provider, poison control center). The question arises as to the reaction to reformulated OxyContin of individuals who abuse prescription opioids and are not reporting abuse to researchers or other authorities. It is of further interest to monitor and describe the extent to which individuals are engaging in efforts to defeat the tamper-resistant properties of reformulated OxyContin and whether such efforts were deemed feasible.

Introduction of reformulated OxyContin presents an opportunity to determine the utility of monitoring Internet data to evaluate reactions to this formulation among a difficult-to-reach population of prescription drug abusers who are not generally in contact with some authority [12]. Because these Internet data reflect uninhibited peer-to-peer communications, they may be a useful source for monitoring and tracking efforts to defeat the abuse-deterrent properties of the product for illicit use. It is generally believed that these efforts will take the form of "recipes" that will be disseminated via the Internet [13-15]. Furthermore, it is anticipated that the feasibility and utility of a recipe will be evaluated by abusers online and that practical tampering methods will be disseminated and perpetuated through postings on websites dedicated to recreational abuse of drugs [15]. Based on this scenario, public health stakeholders are increasingly concerned about monitoring discussions around extraction techniques that emerge on the Internet and tracking the dissemination of these methods [2].

Although public Internet forums can be monitored unobtrusively and might reveal ways in which prescription drugs are being misused [16], there has been little published to date on how to collect, analyze, and understand the messages within the large volume of posts available from online recreational drug abuse communities. Early studies [17,18] that examined the feasibility of systematic Internet surveillance of discussion of prescription opioid products indicated that Internet posts can be reliably coded for sentiment (eg, endorsing vs discouraging abuse) and that both the amount of discussion and sentiment differentiated products [18]. In subsequent work, McNaughton et al [12] developed a metric, referred to as the endorsement ratio (ERo), to evaluate and quantify the overall sentiment expressed by a large number of opioid abusers who post online about prescription opioid products.

In the present work, we sought to understand how drug abusers reacted to the introduction of an intended tamper-resistant prescription opioid product to the market. We examined data from abusers who participated in Internet message boards to evaluate discussion of OxyContin before and after introduction of the reformulation. Specifically, we investigated these questions: (1) did the level of Internet discussion related to OxyContin change quantitatively over time following introduction of the reformulated version of the product, (2) within the OxyContin-specific discussion that did occur, was there a shift in the sentiment expressed by abusers who posted on these websites following the introduction of reformulated OxyContin, and (3) given concerns about efforts to generate and disseminate tampering methods intended to defeat the properties of reformulated OxyContin for use by unintended ROAs, could Internet discussion of such recipes be defined, identified, and monitored?

Methods

Study Overview

The study aimed to evaluate the potential effect the introduction of the reformulation of OxyContin had on discussion within Internet-based recreational drug abuse message boards. Over the pre-post reformulated OxyContin timeframe, we conducted (1) a quantitative evaluation of message board discussion for OxyContin and comparators to capture the relative levels of discussion and any changes during the pre-post time period, (2) a qualitative coding of Internet post content and estimation of endorsement for OxyContin and comparators to determine any changes in the sentiment in favor of each medication for abuse purposes from pre to post OxyContin reformulation, and (3) in the period following the introduction of OxyContin, an evaluation of Internet post content related to tampering methods for defeating the abuse-deterrent properties of reformulated OxyContin. All research activities conducted for this study were exempt from Institutional Review Board review as determined by the New England Institutional Review Board.

For the quantitative evaluation of discussion levels and content analysis/estimation of endorsement, Vicodin (hydrocodone) and Dilaudid (hydromorphone) were selected as comparators. These comparators represented a widely available and highly abused prescription opioid (Vicodin) and a high-potency opioid analgesic that is highly desirable for abuse (Dilaudid) [19]. In order to make appropriate comparison to the target product (OxyContin), qualitative coding and analysis was restricted to discussion of the proprietary products Vicodin and Dilaudid only and did not include generic references to hydrocodone, hydromorphone, and other proprietary products within the opioid compounds (eg, Lortab for hydrocodone and Exalgo for hydromorphone).

Data Source

The study sample consisted of Internet posts (ie, messages) copied from 7 publically accessible message boards that represent a population of drug abusers and their online communications regarding both illicit and prescription drugs. The websites were chosen based upon predefined criteria as described in McNaughton et al [12]. All posts written between January 1, 2008 and September 30, 2013 (N=6,891,514) were archived in a database for further sampling and analysis. No personal identifiable information related to the author was retained.

Quantitative Evaluation of Message Board Discussion

From the database of saved Internet posts, all messages related to OxyContin (both original and reformulated versions of the product), Vicodin, and Dilaudid written between January 1, 2008 and September 30, 2013 (ie, Q1 2008 through Q3 2013) were identified through the use of standardized queries. These queries contained text-matching criteria that included common misspellings, slang, and wildcard characters as well as exclusion criteria to capture as many relevant posts as possible while minimizing the number of false positives (ie, posts returned by the query that are not actually related to the target product) selected. It should be noted, however, that false positives could not be completely eliminated from the text-matching query results without manual review, which was not conducted for this analysis because of the magnitude of posts involved. The rate of discussion related to each product was then calculated as the number of product-specific posts identified per 1000 posts saved within the database per quarter.

Formal Content Analysis and Estimation of Endorsement

A formal content analysis was conducted on random samples of Internet posts related to OxyContin, Vicodin, and Dilaudid during the 26-month period before (preintroduction period=June 1, 2008 through July 30, 2010) and the 26-month period after the introduction of reformulated OxyContin (postintroduction period=August 1, 2010 through September 30, 2012) and identified through the use of the standardized queries. For this analysis, posts retained for coding in the preintroduction period pertained to the original formulation of OxyContin, whereas posts sampled and retained in the postintroduction period related specifically to reformulated OxyContin. Because the design involved comparison of discussion of original OxyContin in the preintroduction period and reformulated OxyContin in the postintroduction period, discussion of original OxyContin in the postintroduction period was not examined for this study. Using systematic query searches, product-specific Internet posts were randomly sampled from the archive. All coding was conducted as part of a larger dynamic postmarketing surveillance program, involving rolling sampling and content analysis of posts (ie, multiple waves of sampling throughout the study period). Power analyses to determine the sample size needed to detect changes were calculated periodically throughout surveillance and changed over time resulting in somewhat different sample sizes in the preintroduction and postintroduction periods for this evaluation.

The coding procedure and assessment of intercoder agreement used in this study is described in detail in McNaughton et al [12]. Briefly, posts were reviewed by trained coders and categorized as either abuse-related or non-abuse-related, and false positives were removed and replaced. A false positive is a query-selected post that upon manual review did not pertain to the specified prescription opioid product. Within the sample of abuse-related posts, product-specific content was further coded as endorsing, discouraging, mixed, or unclear (ie, the sentiment was assigned) (Figure 1). When there was disagreement between coders, the post content was discussed and reviewed by an independent lead coder for a final rating and to achieve a final set of codes for analysis. To assess reliability of the coding, 20% of all posts were coded by 2 coders who were blinded to which posts were coded by both coders and which were coded independently. Interrater agreement (kappa) was then calculated on the 20% overlapping sample to determine if an acceptable level of coder reliability was achieved [20].

A mixed effects multinomial logistic regression was employed to model the probability of observing each of the 4 types of abuse-related Internet posts (endorsing, discouraging, mixed, and unclear) per product. The fixed effects included a product indicator (1=product A, 2=product B, etc), time indicator

(1=preintroduction period, 2=postintroduction period) and product×time interaction. An author random effect was incorporated in the model to account for correlation among messages posted by the same author. The GLIMMIX procedure in SAS 9.3 (SAS Institute, Inc, Cary, NC, USA) was used to fit the model, producing the following statistics of interest:

1. Probability of observing each type of abuse-related post (endorsing, discouraging, mixed, and unclear) per product in the period before and after the introduction of reformulated OxyContin.
2. Endorsement ratio (ERo) for each product in the period before and after the reformulation of OxyContin. The ERo provides a relative estimate of the extent to which a product was being endorsed during each time period by calculating a ratio of probabilities (eg, probability of endorsing product A in the postintroduction period divided by probability of discouraging product A in the postintroduction period), commonly referred to as a relative risk [12].
3. Post-to-preintroduction change in the ERo was estimated by calculating the ratio of EROs (eg, ERo of product A in the postintroduction period divided by ERo of product A in the preintroduction period), commonly referred to as a relative risk ratio.
4. Within-author correlation as estimated by intraclass correlation coefficients derived from the variance components [21].

Figure 1. Abuse-related sentiment categories in formal content analysis.

Sentiment category	Definition and examples
Endorsing	Pertains to posts that endorse, promote, or facilitate use of a compound/product in any way other than how it is prescribed. Examples include: • *I really like product X!* • *If you can't find compound Y, try compound X.* • *I heard you can snort compound X, is this true?* • *If I take product X, what dosage would be good?*
Discouraging	Posts that discuss strong negative opinions in relation to a compound/product. This includes references to addiction, withdrawal, barriers to recreational use, and experiences that an author clearly indicates as negative and or expresses intention not to use the drug again. Examples include: • *I hate compound X, it is the worst drug.* • *You won't get high if you take product X.* • *My friend said that compound X causes the worst withdrawals; I never plan on using it.*
Mixed	Pertains to posts that contain both positive and negative content related to the compound/product. Examples include: • *You shouldn't inject that product because it is dangerous, but if you do it, you can use this method…* • *I think compound X is ok, but not nearly as great as compound Y.* • *I get sick after taking product X sometimes, but it really isn't that bad.*
Unclear	Posts contain unclear, neutral, factual, or indecisive content related to the compound/product. Examples include: • *Compound X is manufactured by Company A.* • *Did you hear the song where product X was referenced?* • *I had a dream the other day…*

Evaluation of Recipes

Of particular interest with respect to any purported ADF product is whether tampering methods, or recipes, are developed that allow individuals to readily defeat the abuse-deterrent properties of a new formulation. To evaluate this possibility, a review of recipe-related content was performed on Internet posts pertaining to reformulated OxyContin during the approximately 3-year period following the product's launch (August 9, 2010 through September 30, 2013). For this evaluation, a recipe was defined as a process (physical, chemical, or potentiation) that enabled use of the product in a way other than intended (ie, swallowing a tablet whole) because ADFs are not formulated to prevent abuse by swallowing multiple tablets whole at one time [4,5]. While variation existed with respect to (1) the format in which a recipe was communicated (eg, step-by-step instruction guide

vs narrative experience report), (2) the words used to describe a recipe, and (3) the devices used by an individual for a particular recipe; recipes were classified into profiles that represented the fundamental or basic steps used when manipulating a product. For example, 2 posts, one that references "crushing a tablet with a knife before putting in water" and a second that notes "(1) pound product with a hammer, (2) add water" would be classified as the same recipe profile (ie, crush and dissolve) despite differences in the format, words, and devices communicated.

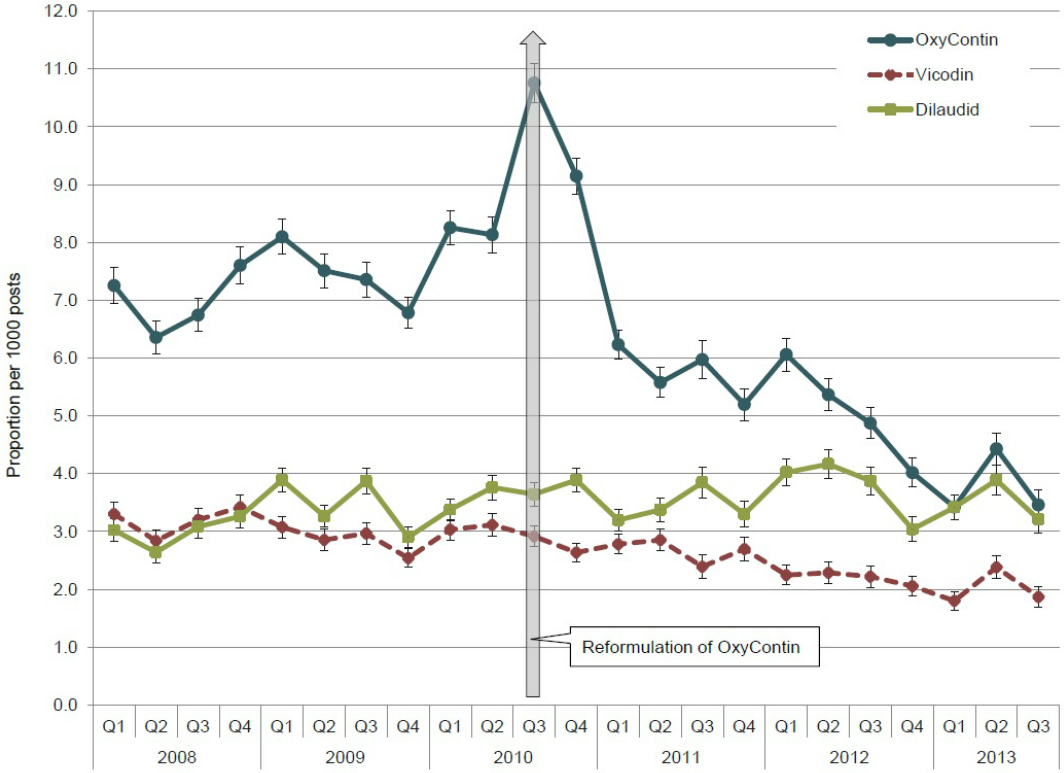
All posts that referenced OxyContin during the approximately 3-year period were reviewed by a trained coder for recipe content related to the reformulated version of OxyContin. For each post that mentioned a recipe related to reformulated OxyContin, the coder assigned 3 codes: (1) the recipe profile, (2) the ROA mentioned in relation to the recipe profile, and (3) whether the author described the recipe as "feasible." Feasibility was defined as being able to manipulate reformulated OxyContin for abuse via an unintended ROA (ie, use of the product other than swallowing the tablet whole). Utilizing the coded information, the total number of recipe-related posts, recipe profiles, and the frequency in which recipe profiles were first observed are presented. In addition, the ROAs mentioned in relation to feasible recipes are provided.

Results

Quantitative Evaluation of Message Board Discussion

Between January 1, 2008 and September 30, 2013 (ie, Q1 2008 through Q3 2013), 45,936 posts related to OxyContin (original formulation in the preintroduction period and both original and reformulated versions of the product in the postintroduction period) were identified in the database of 6,891,514 saved posts. Because the brand name of OxyContin did not change following the introduction of the reformulation, it was not possible to disambiguate references to original versus reformulated OxyContin in the postintroduction period without review of each post, which was not conducted for this analysis. In addition, 18,685 posts related to Vicodin, and 23,863 posts related to Dilaudid were identified. When evaluated by quarter, the proportion of OxyContin-related posts fluctuated between 6.35 and 8.25 posts per 1000 posts during the period before the release of reformulated OxyContin (Q1 2008 through Q2 2010) before increasing to an observed 10.76 posts per 1000 posts in Q3 2010 with the launch of reformulated OxyContin on August 9, 2010 (Figure 2). Following the release of reformulated OxyContin, the proportion of OxyContin posts remained elevated at 9.15 posts per 1000 posts in Q4 2010 before decreasing in Q1 2011 (6.23 posts per 1000). From Q1 2011 through Q3 2013, the proportion of OxyContin-related posts decreased over time, from 6.23 posts per 1000 posts in Q1 2011 to 3.46 posts per 1000 posts in Q3 2013 and remained consistently lower than the quarterly proportions observed before the release of the reformulation (Q2 2008 through Q2 2010). Changes in the proportion of OxyContin-related posts before and following release of the reformulated version of the product, however, contrast with the comparatively consistent pattern of discussion observed for both Vicodin (range 1.87-3.30 posts per 1000 posts) and Dilaudid (range 2.64-4.16 posts per 1000 posts) during the same time period (Figure 2).

Figure 2. Proportion of OxyContin-, Vicodin-, and Dilaudid-related posts from Q1 2008 to Q3 2013.



Formal Content Analysis and Estimation of Endorsement

Coding Results

Of the 16,588 posts sampled, 5365 (32.40%) were identified as false positives and removed from the final sample (Table 1). The high false positive rate for the entire sample was primarily attributable to the number of false positive posts associated with reformulated OxyContin in the postintroduction period. Using the rolling sampling procedures, a total of 11,223 posts were retained for analysis: 3741 posts for each product. For the 20% overlapping sample (ie, posts coded by both reviewers), kappa was calculated on (1) whether a post was abuse-related and, if abuse-related, (2) whether the content was endorsing, discouraging, mixed, or unclear (ie, the sentiment). Kappa statistics were calculated by product and period as well as across all compounds. All kappas were satisfactory (Table 1) [20].

Table 1. Content analysis: number of posts sampled and retained, false positive rate, and interrater agreement kappa statistics.

Product and period[a]	Total posts sampled, n	False positives, n (%)	Final sample, n	Kappa	
				Abuse-related	Sentiment
OxyContin [b]					
Pre	2256	287 (12.72)	1969	.82	.68
Post	5750	3978 (69.18)[c]	1772	.87	.77
Vicodin					
Pre	2226	257 (11.55)	1969	.78	.64
Post	2053	281 (13.69)	1772	.78	.63
Dilaudid					
Pre	2233	264 (11.82)	1969	.83	.68
Post	2070	298 (14.40)	1772	.75	.65
Total	16,558	5365 (32.40)	11,223	.87	.72

[a]Preintroduction period: the period before the reformulation of OxyContin (June 1, 2008 through July 31, 2010); postintroduction period: the period following the reformulation of OxyContin (August 1, 2010 through September 30, 2012).

[b]Preintroduction period represents content related to the original formulation of OxyContin. Postintroduction period represents content related to the reformulated version of OxyContin.

[c]The high false positive rate observed for reformulated OxyContin during the postintroduction period was related to the slang term "OP" (ie, the indicia on the reformulated tablet) which is also an acronym commonly used on message boards to refer to the "original poster" or the first author to write a post in a thread. Furthermore, the standardized queries sometimes yielded posts in which the discussion could not clearly be identified as pertaining to the reformulated version of OxyContin specifically, which resulted in a high degree of false positives.

Estimation of Endorsement

The probability of observing endorsing, discouraging, mixed, or unclear abuse-related sentiments for OxyContin, Vicodin, and Dilaudid in the periods before (preintroduction period) and following (postintroduction period) the introduction of reformulated OxyContin is presented in Table 2. Overall, the probability of observing content related to each sentiment category varied by product and period. For OxyContin specifically, the probability of observing posts with endorsing sentiment was greater for the original formulation (preintroduction period: probability=0.43) than for the reformulated version (postintroduction period: probability=0.22). Conversely, the probability of observing a discouraging post was lower for the original formulation in the preintroduction period (probability=0.22) than for the reformulated version in the postintroduction period (probability=0.27). When evaluated as an ERo [12] as a means of estimating the extent to which the product was endorsed, in the period before the release of reformulated OxyContin, the probability of observing posts that endorsed the use of the original formulation of OxyContin was approximately 1.91 times greater than the probability of discouraging the product (Table 3). In the postintroduction period, however, reformulated OxyContin was 1.23 times more likely to be discouraged than endorsed (ERo=0.81). Taken together, the change in the ERo estimates before and after the introduction of reformulated OxyContin indicate that the ERo for the original formulation of OxyContin in the preintroduction period was 2.33 times greater than the ERo estimate for the reformulated version of OxyContin in the postintroduction period (ratio of EROs=0.43, P<.001) (Table 3).

Table 2. Abuse-related sentiment category probabilities.

Product and period[a]	Endorsing		Discouraging		Mixed		Unclear	
	Prob	95% CI	Prob	95% CI	Prob	95% CI	Prob	95% CI
OxyContin [b]								
Pre	0.43	0.40-0.45	0.22	0.20-0.25	0.18	0.16-0.20	0.16	0.14-0.18
Post	0.22	0.20-0.24	0.27	0.24-0.30	0.23	0.21-0.25	0.28	0.25-0.31
Vicodin								
Pre	0.36	0.33-0.38	0.29	0.26-0.31	0.22	0.20-0.24	0.13	0.12-0.15
Post	0.35	0.32-0.37	0.17	0.15-0.19	0.28	0.26-0.30	0.20	0.18-0.22
Dilaudid								
Pre	0.46	0.43-0.49	0.19	0.17-0.22	0.25	0.23-0.28	0.09	0.08-0.11
Post	0.47	0.44-0.49	0.09	0.08-0.11	0.31	0.28-0.33	0.13	0.11-0.15

[a]Preintroduction period: the period before the reformulation of OxyContin (June 1, 2008 through July 31, 2010); postintroduction period: the period following the reformulation of OxyContin (August 1, 2010 through September 30, 2012).

[b]Preintroduction period represents content related to the original formulation of OxyContin. Postintroduction period represents content related to the reformulated version of OxyContin.

Table 3. Endorsement ratios (ERo) and post-to-preintroduction period ratios of Eros.

Product and period[a]	ERo[b]	95% CI	Ratio of EROs[c]	95% CI	*P*
OxyContin [d]			0.43	0.35-0.52	<.001
Pre	1.91	1.66-2.20			
Post	0.81	0.69-0.95			
Vicodin			1.66	1.36-2.04	<.001
Pre	1.24	1.08-1.42			
Post	2.06	1.76-2.43			
Dilaudid			2.11	1.68-2.63	<.001
Pre	2.38	2.05-2.77			
Post	5.01	4.15-6.05			

[a]Preintroduction period: the period before the reformulation of OxyContin (June 1, 2008 through July 31, 2010); postintroduction period: the period following the reformulation of OxyContin (August 1, 2010 through September 30, 2012).

[b]The ERo is a ratio of probabilities (eg, probability of endorsing product A in the postintroduction period divided by probability of discouraging product A in the postintroduction period), which is commonly referred to as a relative risk.

[c]The post-to-preintroduction ratio of EROs is an estimate of the change the ERo before and after the introduction of reformulated OxyContin (eg, ERo of product A in the postintroduction period divided by ERo of product A preintroduction period), which is commonly referred to as a relative risk ratio.

[d]Preintroduction period represents content related to the original formulation of OxyContin. Postintroduction period represents content related to the reformulated version of OxyContin.

Changes in the sentiment profiles of Vicodin and Dilaudid were also observed before and after the introduction of the reformulated version of OxyContin. In relation to Vicodin, the ERo was 1.66 times greater in the postintroduction period than in the preintroduction period (*P*<.001) indicating that the ratio of encouraging-to-discouraging discussion for Vicodin in the period following the introduction of reformulated OxyContin was significantly greater than in the period before the reformulation (Table 3). Likewise, the ERo for Dilaudid was 2.11 times greater (*P*<.001) in the postintroduction period than in the preintroduction period. These changes in the ERo estimate for Vicodin and Dilaudid, however, appear to be because of a reduction in posts coded as discouraging rather than an increase in encouraging posts. In relation to the post-to-preintroduction period ratio of the EROs for Vicodin and Dilaudid compared to OxyContin, however, the magnitude of the change for Vicodin and Dilaudid was 3.91 times greater (*P*<.001) and 4.95 times greater (*P*<.001) than OxyContin, respectively. The post-to-preintroduction period ratio of the ERo estimates for Vicodin compared to Dilaudid was not statistically different (*P*=.12). These results suggest that the endorsing and discouraging sentiment profile for OxyContin before and after the introduction of the reformulation changed significantly more than the endorsing and discouraging sentiment profile of Vicodin and Dilaudid.

Evaluation of Recipes

During the approximate 3-year period following the launch of reformulated OxyContin (August 9, 2010 through September 30, 2013), 19,659 posts related to OxyContin (both original and reformulated versions of the product) were identified and reviewed by trained coders (Figure 3). Of these, 5677 posts were identified as referring specifically to the reformulated version of OxyContin. Within this reformulated OxyContin-specific discussion, recipes related to reformulated OxyContin were mentioned 1052 times within 825 posts (14.5% of reformulated OxyContin-related discussion) and evidence of feasible manipulation of reformulated OxyContin (ie, use of the product other than swallowing the tablet whole) was observed 576 times within 498 posts (8.8% of reformulated OxyContin-related discussion) across the approximately 3-year period. As Figure 4 illustrates, the frequency of OxyContin-related posts peaked with the introduction of reformulated OxyContin and then declined steadily. Figure 4 also shows a general decrease over the approximately 3-year period in the number of posts specifically referencing a reformulated OxyContin recipe as well as posts that specifically mentioned a feasible recipe. An exception to this general decrease was a slight increase in Q1 2012, which is likely related to discussion associated with the launch of a reformulated version of extended-release oxymorphone. Specifically, authors discussed their experience with reformulated OxyContin recipes and whether or not those methods could be used with the reformulated version of extended-release oxymorphone.

In total, 37 unique recipe profiles were identified during the approximately 3-year period, 32 of which were denoted as feasible at least once (Table 4). Within the reformulated OxyContin recipe-related posts, most referenced 12 of the 37 profiles, whereas the remaining 25 were mentioned fewer than 10 times each during the approximately 3-year period. The frequency with which new recipe profiles emerged decreased following the first quarter after the launch of reformulated OxyContin (from 24 in Q3 2010 [ie, August 9, 2010 to September 30, 2010] to 3 in Q4 2010), and few new recipe profiles were identified in subsequent quarters (Figure 5). Likewise, the number of new feasible reformulated OxyContin recipe profiles observed over time followed a similar pattern.

When considering the 498 reformulated OxyContin-related posts that referenced a feasible recipe profile, various ROAs were mentioned in relation to use of the manipulated product (Figure 3). Oral use of reformulated OxyContin following feasible use of a recipe (eg, drinking in solution, chewing, parachuting) was mentioned in 4.58% (260/5677) of all reformulated OxyContin-related discussion, followed by snorting in 2.25% (128/5677) of reformulated-related discussion, and injection in 0.72% (41/5677) of reformulated OxyContin-related discussion. Smoking or rectal administration of reformulated OxyContin following feasible manipulation were observed 7 and 6 times, respectively, during the approximately 3-year period following the introduction of the reformulated version of OxyContin. It should be noted that an author could reference more than 1 recipe profile as well as more than 1 ROA in relation to a recipe profile within the same post; therefore, the ROA categories within the 498 posts that mentioned feasible recipe profiles are not mutually exclusive. Furthermore, some authors did not indicate use of a specific ROA (99/5677, 1.74%) following feasible manipulation.

Table 4. Frequency of reformulated OxyContin recipe profiles.

Recipe profile	Posts that mentioned recipe profile, n	Posts that mentioned recipe profile was feasible, n
Crush/shave	277	152
Dissolve/soak	130	58
Chew	114	81
Crush/shave, heat, and freeze	89	72
Crush/shave and dissolve/soak	71	39
Crisp	50	35
Crush/shave and heat	40	26
Crush/shave, heat, and dissolve/soak	25	17
Crush/shave, add chemicals, and evaporate	24	13
Take with acidic foods or beverages	23	19
Heat	19	6
Take with alcohol	10	6
Dissolve/soak and heat	9	4
Heat and freeze	8	2
Crush/shave, heat, freeze, and dissolve/soak	7	7
Crush/shave and freeze (or vice versa)	7	5
Crush/shave, add chemicals, dissolve, and filter	5	4
Take with a fatty meal	5	4
Crush/shave, heat, cool, dissolve/soak, and filter	4	3
Crush/shave, heat, dissolve/soak, and evaporate	4	3
Dissolve/soak and filter	4	1
Add soda, heat, and take with acidic beverage	4	3
Crisp, dissolve/soak, and filter	4	1
Dissolve/soak and freeze	3	0
Crush/shave, heat, dissolve/soak in chemical	3	0
Crush/shave, dissolve/soak, and evaporate	2	0
Crush/shave, heat, dissolve/soak, and filter	2	2
Heat, cool, crush/shave, dissolve/soak, and heat	2	3
Freeze, crush/shave, heat, and crush/shave	2	3
Crisp, heat, and freeze	2	1
Crush/shave, add chemicals, evaporate, and heat	2	1
Crush/shave, dissolve/soak, and cool/freeze	1	0
Heat, crush/shave, heat, and freeze	1	1
Crush/shave, add chemicals, evaporate, and cool	1	1
Dissolve/soak, heat, and filter	1	0
Crush/shave, heat, freeze, crisp, and filter	1	1
Crush/shave, dissolve/soak, add chemical	1	1
Total	825[a]	498[a]

[a]An author could reference more than 1 recipe profile within the same post; therefore, the total number of recipe-related posts does not equal the sum of the counts across the 37 recipe profiles.

Figure 3. Evaluation of recipes: coding results. Feasibility was defined as being able to manipulate reformulated OxyContin for abuse via an unintended route of administration (ie, use of product other than swallowing the tablet whole).

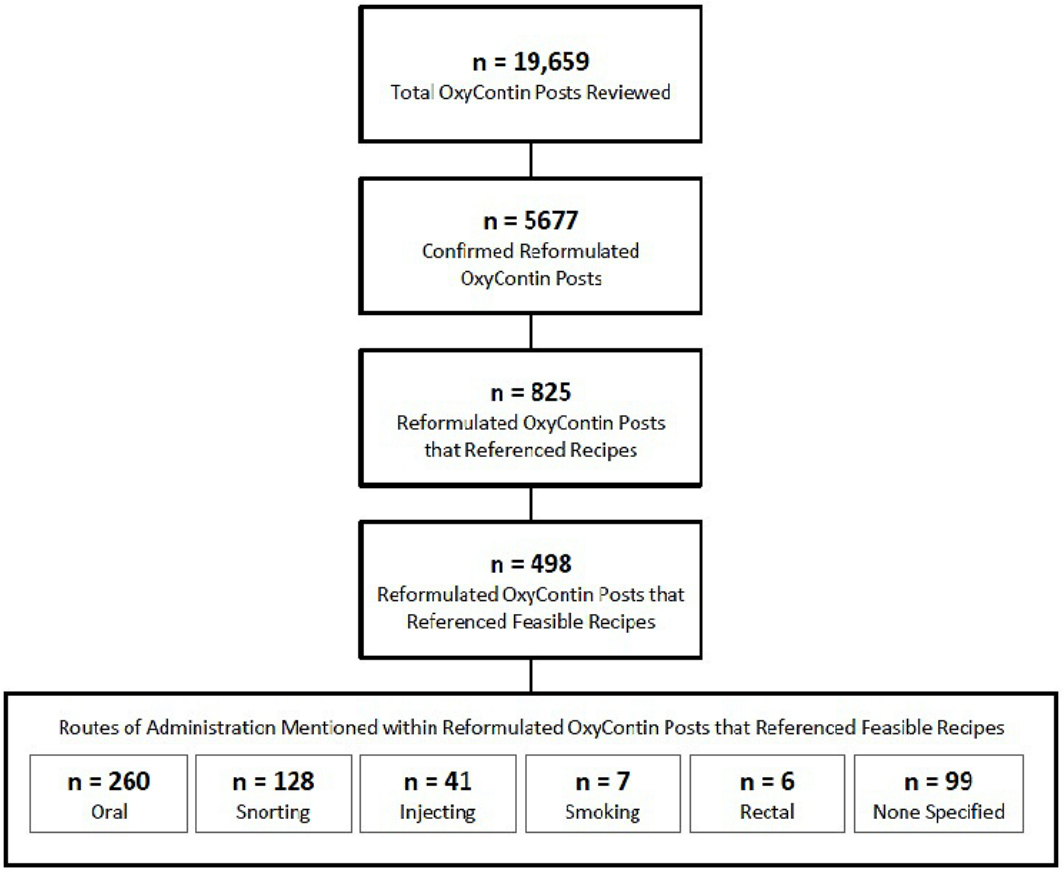


Figure 4. Frequencies of OxyContin-, reformulated OxyContin recipe-, and feasible reformulated OxyContin recipe-related posts from Q3 2009 to Q3 2013. For the reformulated OxyContin recipe and feasible OxyContin recipe categories, Q3 2010 includes data from August 9, 2010 to September 30, 2010.

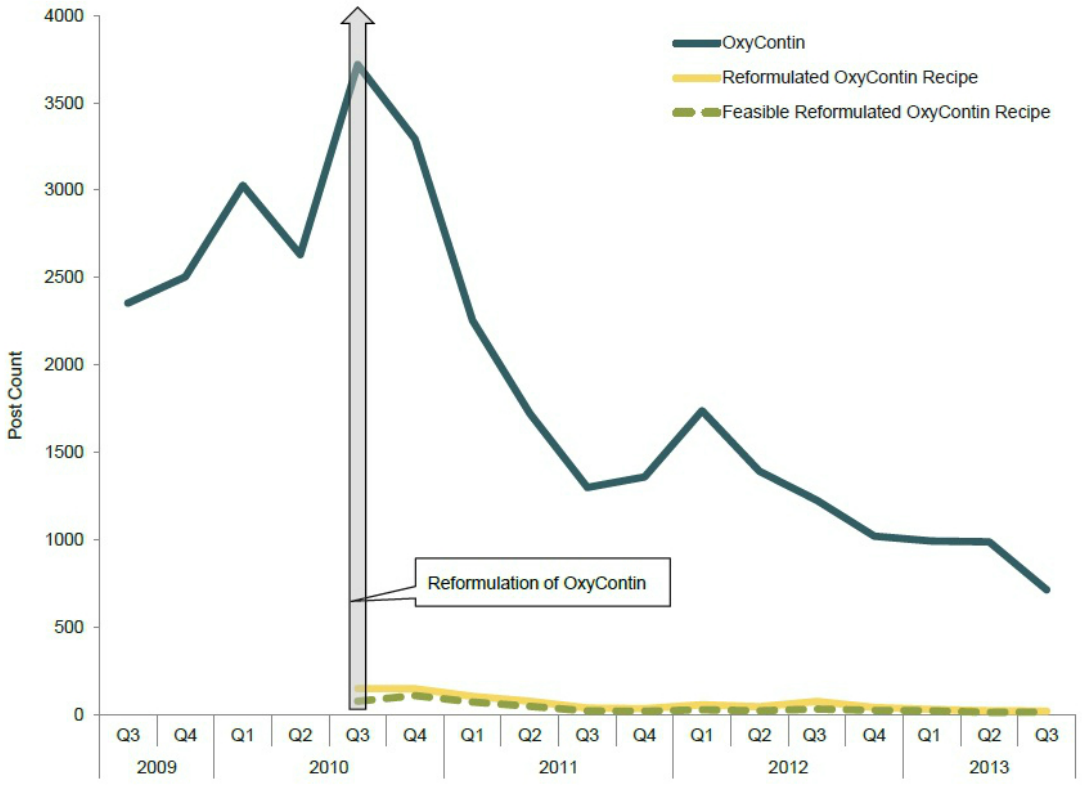
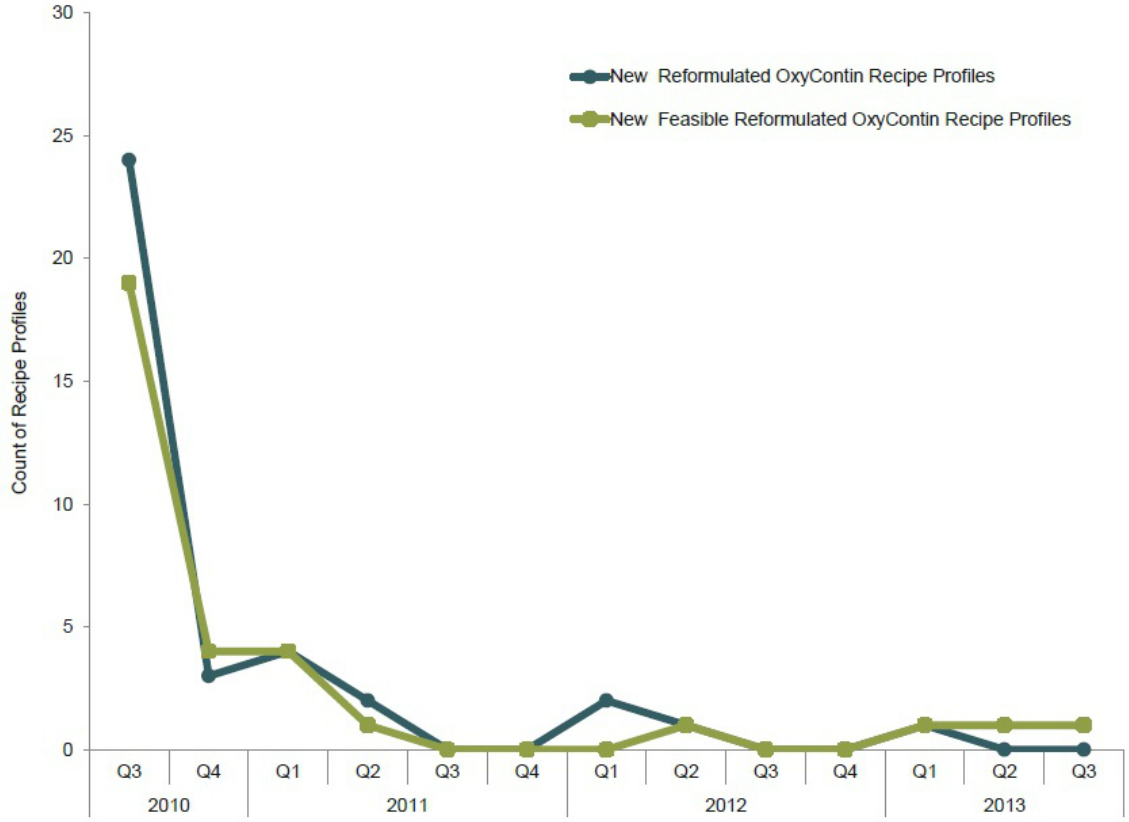


Figure 5. Frequency of new reformulated OxyContin recipe profiles and new feasible reformulated OxyContin recipe profiles.



Discussion

Principal Findings

Novel methodologies for evaluating the impact of prescription opioid products with abuse-deterrent properties on abuse-related behaviors are needed. This study presents an Internet-focused approach to examine discussion on recreational drug abuse Internet forums regarding reformulated OxyContin, the first widely available reformulated opioid product on the market. Systematic monitoring and review of content from online message boards before and after the introduction of reformulated OxyContin allowed for evaluation of (1) discussion about OxyContin and 2 comparators, (2) relative endorsement of OxyContin and comparators, and (3) discussions around manipulation of reformulated OxyContin through recipes.

Overall, the findings presented here suggest that the introduction of reformulated OxyContin had an impact on discussion of OxyContin on message boards frequented by prescription drug abusers. Quantitative analysis indicated that the volume of discussion related to OxyContin increased in the quarters leading up to the launch of the reformulated version of the product and subsequently decreased to levels lower than those observed in the period before the reformulation. In contrast, the level of discussion associated with the selected comparators (Vicodin, a widely available and highly abused prescription opioid, and Dilaudid, a high-potency opioid analgesic that is highly desirable for abuse [19]) remained consistent across the pre-postintroduction period. Content analyses revealed that sentiment related to OxyContin on the message boards changed after the introduction of the reformulation as reflected by a significant decline in the ERo for OxyContin following the reformulation of the product. That is, the online consensus regarding the desirability of OxyContin for abuse appears to have shifted from a positive sentiment to a relatively and significantly more negative view. Individuals who participated on the message boards reviewed during the study period expressed preference for the original version of OxyContin over the reformulated product as evidenced by both the shift in the sentiment profile and the overall decrease in the level of discussion associated with OxyContin over time during the postintroduction period.

The analyses of sentiment built upon prior work [12] by applying the endorsement ratio methodology; that is, using the ERo to quantify change in sentiment expressed by recreational drug abusers about a product with tamper-resistant properties (reformulated OxyContin) compared with sentiment expressed for the parent product (original formulation OxyContin). Our observations are also consistent with findings from other studies [8-11], including one of a sentinel surveillance sample of individuals assessed for substance use problems in the first 20 months after the introduction of reformulated OxyContin that found that the reformulation impacted abuse patterns of OxyContin [8].

Prior to the launch of any reformulated product, like reformulated OxyContin, concern existed that the product would be greeted with numerous attempts by abusers to defeat the product's tamper-resistant mechanism. A particular concern was that any truly successful recipe would be widely and rapidly disseminated online [15]. Results of the systematic examination presented here suggest that abusers who posted online responded

to the introduction of the reformulation with discussion related to recipes for manipulating reformulated OxyContin for abuse during the first few quarters following the product's launch. However, rather than an increasing level of discussion, we observed a small number of posts related to and mentions of such recipes and a decrease to even smaller numbers over time.

Strengths and Limitations

Findings from this study should be considered in light of its limitations. Querying Internet posts based on selected keywords is incomplete and does not identify all discussion potentially related to a particular topic. Although the methodology described here has the advantage of providing a systematic and consistent approach over time, it is possible that some discussion associated with OxyContin was missed in this analysis. For example, discussion containing references to the product via terms such as "it," "that drug," or "what Joe is using," in which an individual is making an inference or reference within a conversation may have been missed. This may have introduced selection bias in the sample of posts used. However, it seems unlikely that such bias would result in having completely missed or underestimated significant discussion or topics related to the introduction of reformulated OxyContin and the potential change in OxyContin-related discussion over time.

For the formal content analysis and estimation of endorsement, a high false positive rate was observed for reformulated OxyContin during the postintroduction period (ie, the period following the product's reformulation), which was primarily attributable to 2 issues. The slang term "OP" (ie, the indicia on the reformulated OxyContin tablet) is also an acronym commonly used on message boards to refer to the "original poster" or the first author to write a post in a thread. Furthermore, because the brand name of the product did not change following reformulation, the search-string queries often yielded posts that, even with human review (which was conducted for this analysis), could not be clearly identified as pertaining specifically to reformulated OxyContin. Although both of these factors contributed to the high degree of false positives, removal of the non–OxyContin-related content as well as ambiguous references to OxyContin that could not be verified as related to the reformulated version of the product ensured that the sample of posts included for analysis in the postintroduction period reflected the target product (ie, the reformulated version) and would, therefore, minimize the effect of misclassification on the results.

It should be noted that references to feasible recipes in this study refer to an author reporting that he/she was able to manipulate reformulated OxyContin and then use it for recreational purposes. Such reports cannot be verified (ie, someone claiming to have tried a recipe may not be telling the truth). However, individuals who participate in the examined forums represent stable communities of drug users and are self-policing so that posted information that is inconsistent with others' experience tends to be "corrected" by the online community. Additionally, reports of feasible recipe use do not necessarily indicate that the desired effect was achieved as a result of the manipulation. Claims of having abused a manipulated product, whether by an oral or a nonoral route (eg, snorting, injecting), does not mean that the effects were equivalent to, better than, or worse than use of the original product. Although one might expect that the overall poorer sentiment observed for reformulated OxyContin suggests dissatisfaction, the present study did not directly examine satisfaction with results of tampering.

Strengths of this study should be highlighted and include (1) the duration of the study period allowed for a large sample size and examination of trends over time, (2) systematic coding of posts with acceptable interrater reliability, (3) the use of operational definitions of recipe profiles that established a standardized methodology for evaluation of Internet content, (4) the ongoing archiving and storage of Internet posts over time allowed the retrospective evaluation of data and avoided bias introduced by forum moderators deleting older posts for reasons of their own (eg, storage space), and (5) the integration of quantitative (number of Internet posts), content (sentiment of individuals), and qualitative analyses (manipulation recipes) provided a comprehensive approach to understanding the reactions of recreational abusers to the introduction of a tamper-resistant product.

Conclusions

This study illustrates the value of analyzing Internet discussion on recreational drug use forums to evaluate the impact of introducing a possible tamper-resistant opioid formulation. Introduction of reformulated OxyContin into the marketplace correlated with changes in discussion of abuse-related behavior among recreational abusers as reflected by changes in online conversation levels, reversal of sentiment about the product, and emergence of manipulation-attempt recipes, consistent with findings from other studies showing reductions in abuse and diversion [8-11]. These findings suggest a possible abuse-deterrent effect of the reformulated product relative to the original formulation that was not observed in comparators.

Acknowledgments

Christine M Benoit, Taryn M Dailey, James Dickinson, and Elizabeth Whalley are acknowledged for their contribution to Internet monitoring activities at Inflexxion, Inc. Preparation of this paper was supported in part by Inflexxion, Inc, Newton, MA, and Purdue Pharma LP, Stamford, CT. The authors had sole editorial rights over the manuscript. This work was supported in part by NIH grant #DA020279.

Conflicts of Interest

Emily C McNaughton, Ryan A Black, Sarah E Weber, and Stephen F Butler are employees of Inflexxion. Inc. Stephen F Butler is also a shareholder of Inflexxion, Inc. Paul M Coplan and Howard D Chilcoat are employees of Purdue Pharma LP.

References

1. Office of National Drug Control Policy. Epidemic: Responding to America's Prescription Drug Crisis. Washington, DC: Executive Office of the President of the United States; 2011. URL: http://www.whitehouse.gov/sites/default/files/ondcp/policy-and-research/rx_abuse_plan.pdf [accessed 2013-02-08] [WebCite Cache ID 6O3696g2Q]
2. US Food and Drug Administration. Transcript for the October 21 Joint Meeting of the Anesthetic and Life Support Drugs Advisory Committee and Drug Safety and Risk Managment Advisory Committee. 2010. URL: http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/AnestheticAndAnalgesicDrugProductsAdvisoryCommittee/UCM236243.pdf [accessed 2013-02-08] [WebCite Cache ID 6O38A7lcc]
3. Katz NP, Adams EH, Benneyan JC, Birnbaum HG, Budman SH, Buzzeo RW, et al. Foundations of opioid risk management. Clin J Pain 2007 Feb;23(2):103-118. [doi: 10.1097/01.ajp.0000210953.86255.8f] [Medline: 17237659]
4. Budman SH, Grimes Serrano JM, Butler SF. Can abuse deterrent formulations make a difference? Expectation and speculation. Harm Reduct J 2009;6:8 [FREE Full text] [doi: 10.1186/1477-7517-6-8] [Medline: 19480676]
5. Katz N. Abuse-deterrent opioid formulations: are they a pipe dream? Curr Rheumatol Rep 2008 Jan;10(1):11-18. [Medline: 18457606]
6. Raffa RB, Pergolizzi JV. Opioid formulations designed to resist/deter abuse. Drugs 2010 Sep 10;70(13):1657-1675. [doi: 10.2165/11537940-000000000-00000] [Medline: 20731474]
7. US Food and Drug Administration. Guidance for Industry: Abuse-Deterrent Opioids - Evaluation and Labeling: Draft Guidance. 2013 Jan. URL: http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM334743.pdf [accessed 2013-02-08] [WebCite Cache ID 6O38Bf0pg]
8. Butler SF, Cassidy TA, Chilcoat H, Black RA, Landau C, Budman SH, et al. Abuse rates and routes of administration of reformulated extended-release oxycodone: initial findings from a sentinel surveillance sample of individuals assessed for substance abuse treatment. J Pain 2013 Apr;14(4):351-358. [doi: 10.1016/j.jpain.2012.08.008] [Medline: 23127293]
9. Cicero TJ, Ellis MS, Surratt HL. Effect of abuse-deterrent formulation of OxyContin. N Engl J Med 2012 Jul 12;367(2):187-189. [doi: 10.1056/NEJMc1204141] [Medline: 22784140]
10. Coplan PM, Kale H, Sandstrom L, Landau C, Chilcoat HD. Changes in oxycodone and heroin exposures in the National Poison Data System after introduction of extended-release oxycodone with abuse-deterrent characteristics. Pharmacoepidemiol Drug Saf 2013 Dec;22(12):1274-1282. [doi: 10.1002/pds.3522] [Medline: 24123484]
11. Severtson SG, Bartelson BB, Davis JM, Muñoz A, Schneider MF, Chilcoat H, et al. Reduced abuse, therapeutic errors, and diversion following reformulation of extended-release oxycodone in 2010. J Pain 2013 Oct;14(10):1122-1130. [doi: 10.1016/j.jpain.2013.04.011] [Medline: 23816949]
12. McNaughton EC, Black RA, Zulueta MG, Budman SH, Butler SF. Measuring online endorsement of prescription opioids abuse: an integrative methodology. Pharmacoepidemiol Drug Saf 2012 Oct;21(10):1081-1092. [doi: 10.1002/pds.3307] [Medline: 22777908]
13. Boyer EW, Shannon M, Hibberd PL. The Internet and psychoactive substance use among innovative drug users. Pediatrics 2005 Feb;115(2):302-305. [doi: 10.1542/peds.2004-1199] [Medline: 15687436]
14. The National Center on Addiction and Substance Abuse at Columbia University. "You've Got Drugs !" Prescription Drug Pushers on the Internet 2006 Update. 2006. URL: http://www.casacolumbia.org/download/file/fid/556 [accessed 2013-02-08] [WebCite Cache ID 6O38DWppH]
15. Cone EJ. Ephemeral profiles of prescription drug and formulation tampering: evolving pseudoscience on the Internet. Drug Alcohol Depend 2006 Jun;83 Suppl 1:S31-S39. [doi: 10.1016/j.drugalcdep.2005.11.027] [Medline: 16458455]
16. Wax PM. Just a Click Away: Recreational Drug Web Sites on the Internet. PEDIATRICS 2002 Jun 01;109(6):e96-e96. [doi: 10.1542/peds.109.6.e96] [Medline: 12042590]
17. Butler S, Katz N, Venuti S, Benoit C, Beaulaurier R. Content analysis and monitoring of Internet prescription opioid abuse-related postings. The Journal of Pain 2006 Apr;7(4):S86-S28. [doi: 10.1016/j.jpain.2006.01.350]
18. Butler SF, Venuti SW, Benoit C, Beaulaurier RL, Houle B, Katz N. Internet surveillance: content analysis and monitoring of product-specific internet prescription opioid abuse-related postings. Clin J Pain 2007 Sep;23(7):619-628. [doi: 10.1097/AJP.0b013e318125c5cf] [Medline: 17710013]
19. Butler SF, Fernandez KC, Chang A, Benoit C, Morey LC, Black R, et al. Measuring attractiveness for abuse of prescription opioids. Pain Med 2010 Jan;11(1):67-80. [doi: 10.1111/j.1526-4637.2009.00736.x] [Medline: 20002325]
20. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174. [Medline: 843571]
21. Snijders TAB, Bosker RJ. Multilevel Analysisan Itroduction to Basic and Advanced Multilevel Modeling. London: Sage Publishers; 2012.

Abbreviations

ADF: abuse-deterrent formulation
ERo: endorsement ratio
ROA: route of administration
TRF: tamper-resistant formulation

Edited by G Eysenbach; submitted 13.03.14; peer-reviewed by S Calderon, J Zacny; comments to author 03.04.14; revised version received 10.04.14; accepted 13.04.14; published 02.05.14.

Please cite as:
McNaughton EC, Coplan PM, Black RA, Weber SE, Chilcoat HD, Butler SF
Monitoring of Internet Forums to Evaluate Reactions to the Introduction of Reformulated OxyContin to Deter Abuse
J Med Internet Res 2014;16(5):e119
URL: http://www.jmir.org/2014/5/e119/
doi:10.2196/jmir.3397
PMID:24800858

Original Paper

Development of a National Agreement on Human Papillomavirus Vaccination in Japan: An Infodemiology Study

Haruka Nakada[1], BS, JD; Koichiro Yuji[1], MD, PhD; Masaharu Tsubokura[1], MD; Yukio Ohsawa[2], PhD; Masahiro Kami[1], MD, PhD

[1]Institute of Medical Science, Division of Social Communication System for Advanced Clinical Research, The University of Tokyo, Tokyo, Japan
[2]School of Engineering, Department of Systems Innovation, The University of Tokyo, Tokyo, Japan

Corresponding Author:
Haruka Nakada, BS, JD
Institute of Medical Science
Division of Social Communication System for Advanced Clinical Research
The University of Tokyo
4-6-1 Shirokanedai,
Minato-ku
Tokyo, 108-8639
Japan
Phone: 81 3 6409 2068
Fax: 81 3 6409 2069
Email: nakurah-tky@umin.net

Abstract

Background: A national agreement on human papillomavirus (HPV) vaccination was achieved relatively quickly in Japan as compared to the United States and India.

Objective: The objective was to identify the role of print and online media references, including references to celebrities or other informants, as factors potentially responsible for the relatively rapid national acceptance of HPV vaccination in Japan.

Methods: A method of text mining was performed to select keywords, representing the context of the target documents, from articles relevant to the promotion of HPV vaccination appearing in major Japanese newspapers and Web pages between January 2009 and July 2010. The selected keywords were classified as positive, negative, or neutral, and the transition of the frequency of their appearance was analyzed.

Results: The number of positive and neutral keywords appearing in newspaper articles increased sharply in early 2010 while the number of negative keywords remained low. The numbers of positive, neutral, and negative keywords appearing in Web pages increased gradually and did not significantly differ by category. Neutral keywords, such as “vaccine” and “prevention,” appeared more frequently in newspaper articles, whereas negative keywords, such as “infertility” and “side effect,” appeared more frequently in Web pages. The extraction of the positive keyword “signature campaign” suggests that vaccine beneficiaries cooperated with providers in promoting HPV vaccination.

Conclusions: The rapid development of a national agreement regarding HPV vaccination in Japan may be primarily attributed to the advocacy of vaccine beneficiaries, supported by advocacy by celebrities and positive reporting by print and online media.

(J Med Internet Res 2014;16(5):e129) doi:10.2196/jmir.2846

KEYWORDS

cervical cancer; health policy; human papillomavirus; public health; vaccination

Introduction

As of October 2010, 29 countries had issued formal recommendations or developed financing plans for the quadrivalent vaccine (Gardasil; Merck) and/or the bivalent vaccine (Cervarix; GlaxoSmithKline) to prevent human papillomavirus (HPV) infection [1]. Many stakeholders likely contributed to the establishment of this national agreement regarding HPV vaccination, including legislators, professional and advocacy organizations, pharmaceutical companies, print and online media, researchers, and informal networks. Of these sources, online media play an increasingly important role in the

sharing of medical information [2]. Popular celebrities' experiences have a tremendous role in raising the public awareness about this disease. In the United Kingdom, cervical cancer screenings increased after a popular TV star, Jade Goody, died from cervical cancer. This is the so-called "Jade Goody effect" [3].

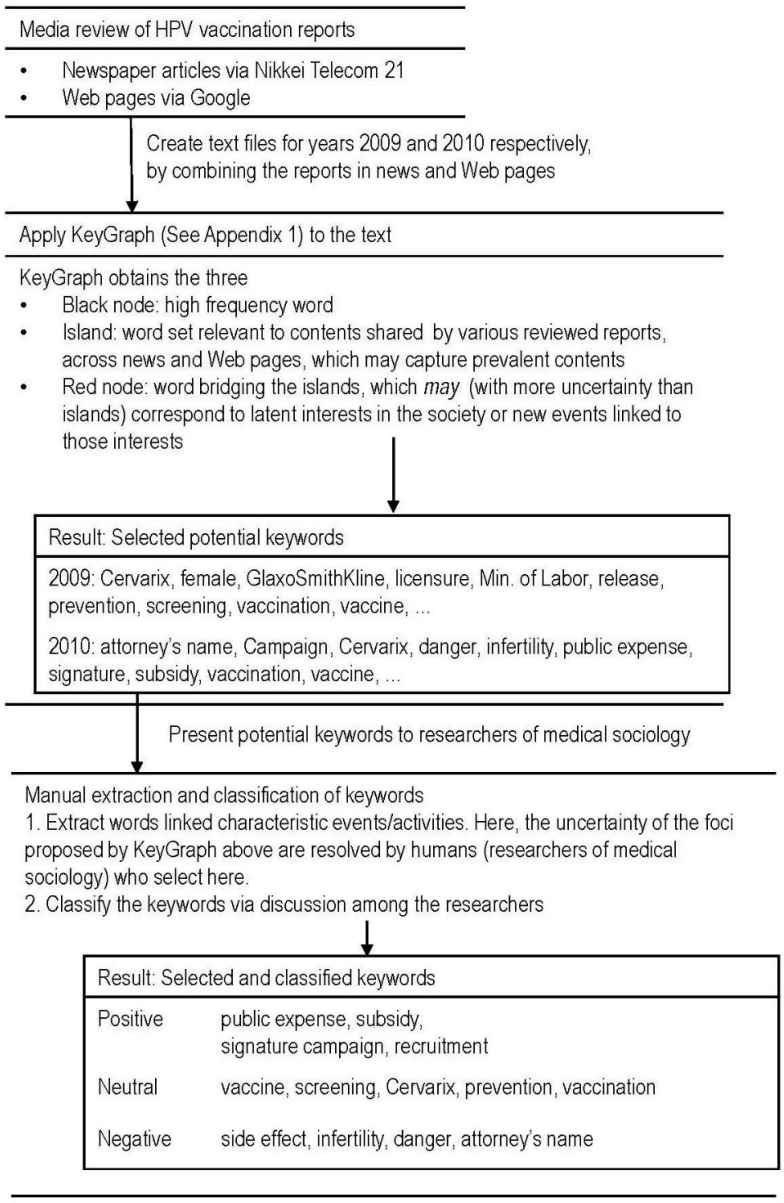
Japan legally supports routine childhood immunization against only 8 diseases [4], a policy that can be primarily attributed to law suits during the 1980s and 1990s [5], and the budget deficit. Nevertheless, unlike in the United States [6] and India [7], national agreement on HPV vaccination was reached relatively smoothly in Japan. After the 2006 licensure of Merck's HPV vaccine in the United States, bills to make HPV vaccination compulsory were introduced in 24 states. However, policymakers changed their mind and as of February 2010, only 2 states had enacted mandates [6]. In India, the government suspended the HPV vaccine trials responding to demands from advocacy groups. In contrast, in Japan, after the licensing of a bivalent HPV vaccine in October 2009, 294 local governments corresponding to 16.4% of the total local governments decided to offer subsidies by October 26, 2010. On August 5, 2010, the Minister of Health, Labor and Welfare proclaimed that they would budget for the national vaccination for the next fiscal year. These actions by the government imply that a national agreement for HPV vaccination was reached. The objective of this study is to identify the role of print and online media references, including references to celebrities or other informants, as factors potentially responsible for the relatively rapid development of the national agreement on HPV vaccination in Japan.

Methods

Overview

An overview of the process and the analysis method used in this paper is shown in Figure 1. We will describe the details of each step in the process, corresponding to each frame in Figure 1, subsequently.

Figure 1. Overview of the process to analyze media review reports on HPV in Japan.



Media Review

After review of articles related to cervical cancer using keywords, a total of 5 newspapers covering approximately 60% of total circulation were selected from among all registered newspapers in Nikkei Telecom 21, the major newspaper database in Japan. Using the Google search engine, we selected Web pages by keywords and the names of 5 Japanese celebrities who survived cervical cancer before 2010; most of these celebrities were involved in activities to promote HPV vaccination or had shared their stories via books or mass media. The monthly change in the number of instances in which each category of keywords appeared in newspaper articles and Web pages was determined and analyzed. We collected Web pages updated for specified periods via Google's search engine. The dates of publication for news articles were listed in the database. All articles and almost all the Web pages were in Japanese because we searched using Japanese keywords (the keywords shown subsequently are translated into English). As we described the situation in Japan, the language was constrainted to Japanese.

Keyword Selection and Classification

Keywords were selected using KeyGraph [8-13] (see Multimedia Appendix 1 for details) and were manually classified into positive, negative, or neutral expressions so that their frequency in newspaper articles and Web pages between January 2009 and July 2010 could be compared.

We used the keyword graphs in Figures 2 and 3 obtained by KeyGraph applied to the review of the reporting of HPV vaccination during the study period. Referring to Multimedia Appendix 1, the number of nodes (element of the document) were set as N_blacknodes=30, N_black links =20, and N_red nodes=10 (the default values for setting these parameters). The black node, red node, and black link are defined as a high frequency word, an important word not shown with high frequency in the document, and a connection in the document, respectively.

From the keywords presented by KeyGraph, 2 domain experts, who are researchers of medical sociology (HN and MK), independently selected words representing already existing concepts relevant to the promotion of HPV vaccine and classified the keywords as positive, neutral, or negative (Table 1), based on discussion with specialists in vaccination policy in Japan. If there was discordance between them, HN made a final decision. The kappa statistic was used to assess the level of agreement among the researchers. The definition of positive, negative, and neutral were "promotion of HPV vaccination," "objection or barrier to HPV vaccination," and "activities against cervical cancer," respectively. For example, "public expense" was positive because HPV vaccination was supported by public funding. An attorney's name was negative because the attorney actively developed a campaign against HPV vaccination.

Table 1. Classification of keywords associated with human papillomavirus (HPV) in Japan.

Category	Keywords
Positive	Public expense, subsidy, signature campaign, recruitment
Neutral	Vaccine, screening, Cervarix, prevention, vaccination
Negative	Side effect, infertility, danger, [an attorney's name]

Figure 2. Keyword map (KeyGraph) showing the primary stakeholders in HPV vaccination in 2009. The most frequent words are shown with black nodespairs of these frequent words co-occurring the most get linked via black lines. "GAN" is a Japanese word meaning cancer.

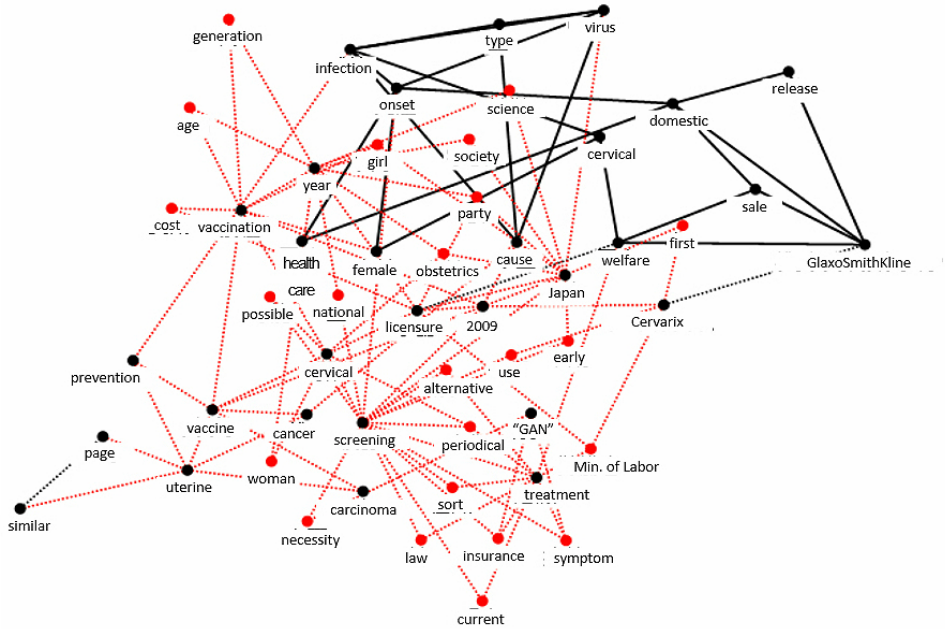
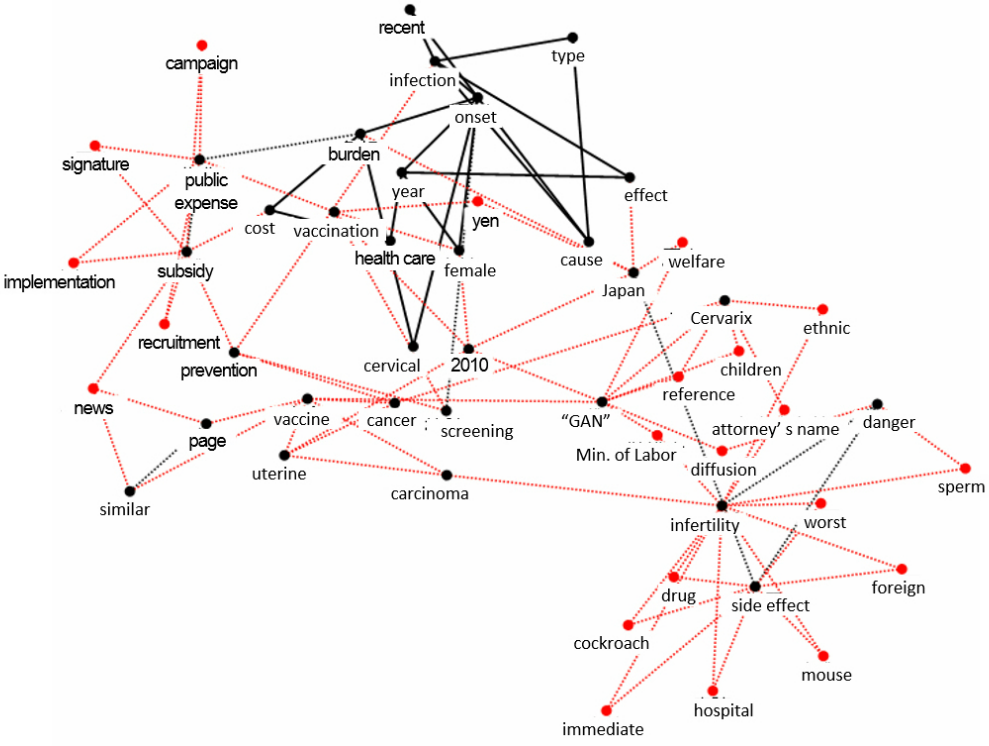


Figure 3. Keyword map (KeyGraph) showing the primary stakeholders in HPV vaccination in 2010. The most frequent words are shown with black nodespairs of these frequent words co-occurring the most get linked via black lines. “GAN” is a Japanese word meaning cancer.



Results

Media Analysis

The total number of newspaper articles and Web pages containing information regarding cervical cancer was 624 and 15,792,600, respectively, for the study period. An increase in the number of newspaper articles coincided with an increase in Web pages in early 2010 (Figures 4 and 5). The numbers of positive and neutral keywords in the newspaper articles increased sharply in early 2010 and peaked in June 2010, whereas the number of negative keywords remained low throughout the study period (Figure 4). In contrast, the numbers of positive, neutral, and negative keywords in Web pages increased gradually over the study period and did not significantly differ by category (Figure 5). Although the number of Web pages containing the names of relevant Japanese celebrities had increased to 19,200 by May 2010 (Figure 4), few newspaper articles reported on these celebrities during the study period (Figure 5). The neutral keywords “vaccine” and “prevention” appeared more frequently in newspapers, whereas the negative keywords “infertility” and “vaccine” appeared more frequently on Web pages.

Figure 4. Serial changes in the number of newspaper articles containing positive, negative, and neutral keywords related to cervical cancer and the names of relevant Japanese celebrities.

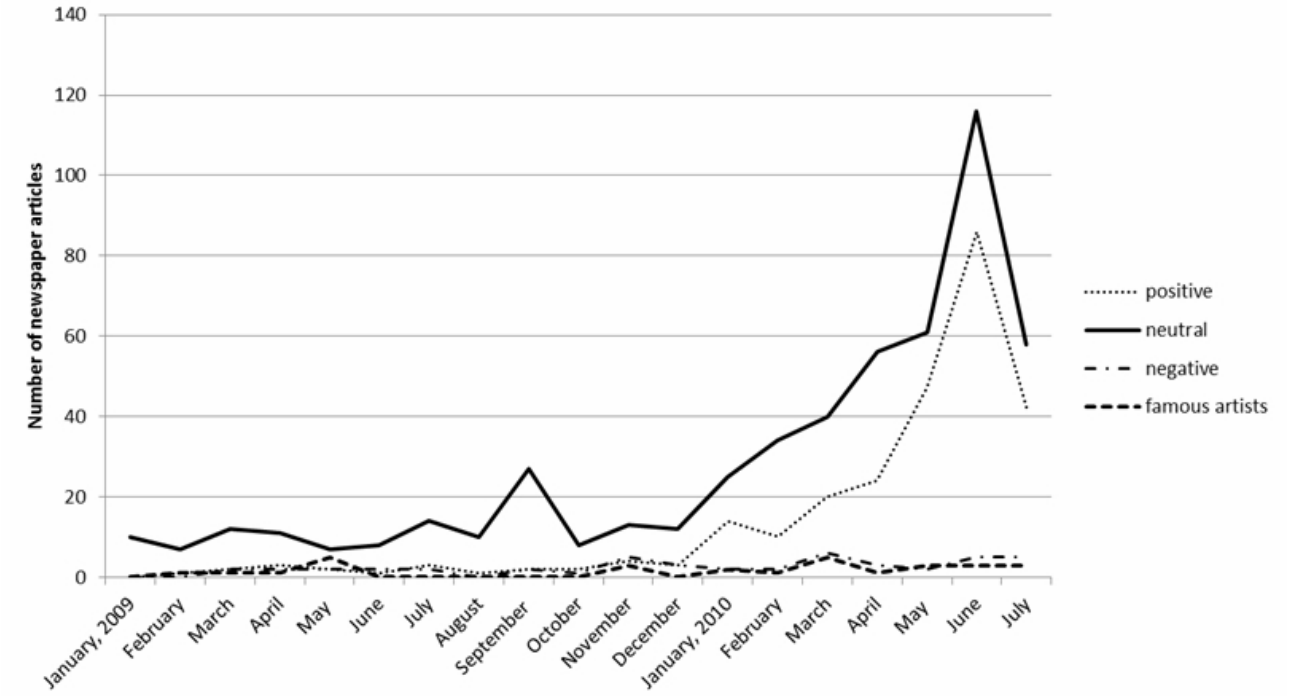
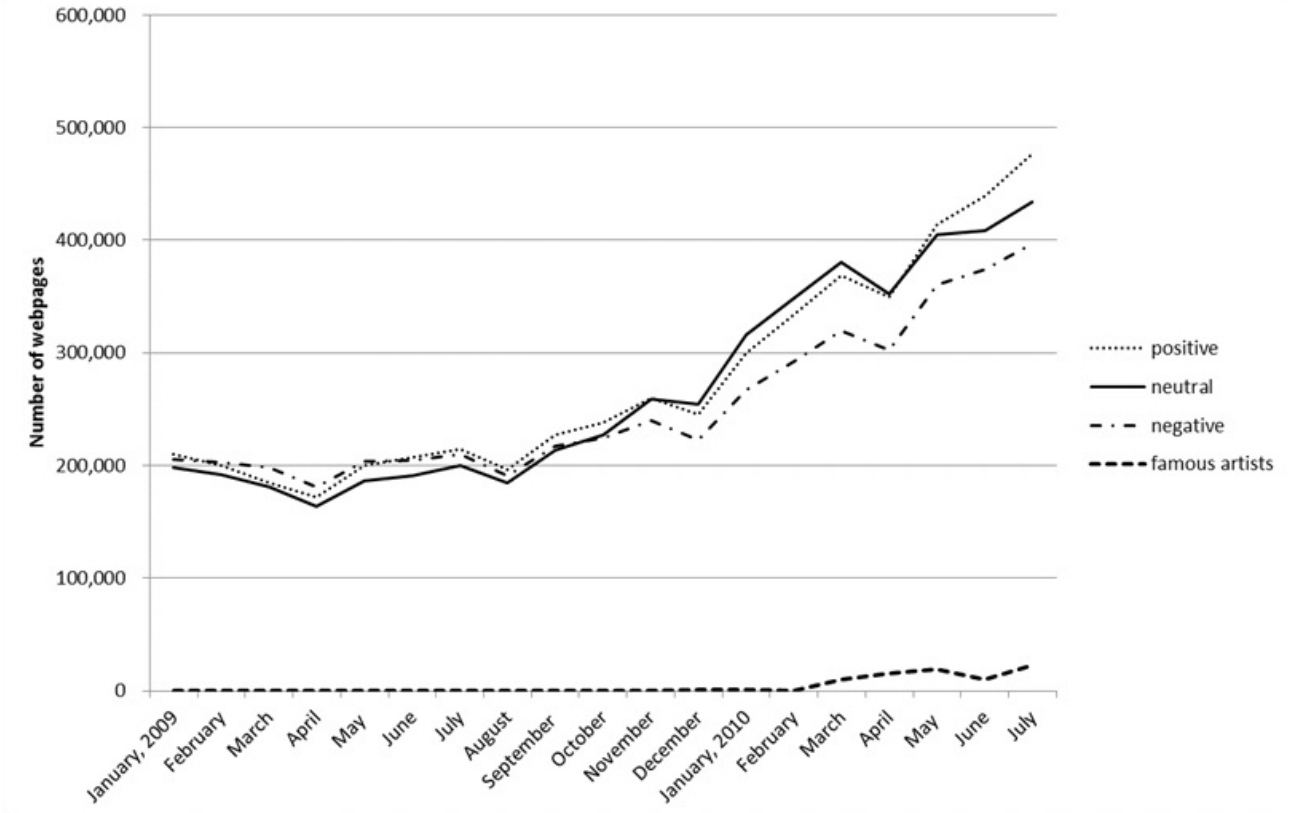


Figure 5. Serial changes in the number of Web pages containing positive, negative, and neutral keywords related to cervical cancer and the names of relevant Japanese celebrities.



Keyword Analysis

The assessment of the agreement on the keyword labeling revealed no significant differences between the 2 researchers (κ=.884). The number of Web pages containing the 3 positive keywords increased gradually over the study period, whereas the number of newspaper articles containing the positive keywords “subsidy” and “public expense” peaked in June 2010. The neutral keywords “vaccine” and “prevention” co-occurred frequently in newspaper articles throughout the study period, with the number containing “prevention” peaking in June 2010. Although the number of Web pages containing the negative

keywords “side effect” and “infertility” increased gradually over the study period, the number of newspaper articles containing “side effect” peaked in June 2010.

Discussion

Principal Findings

Text mining using KeyGraph led to extraction of the keywords “public expense,” “subsidy,” and “signature campaign,” but not “pharmaceutical company” (a main stakeholder in the promotion of HPV vaccination [14-16]) as a positive keyword. The extraction of these words—particularly “signature campaign,” a type of campaign conducted by vaccine beneficiaries—suggests that vaccine beneficiaries cooperated with providers in promoting HPV vaccination in Japan. Because previous studies on HPV vaccination failed to find evidence of such cooperation [17], these findings were unexpected but noteworthy in discussing latent social trends affecting the prevalence of HPV vaccination.

Analysis of the results suggested that the most influential stakeholders had changed from vaccine providers, including medical specialists and pharmaceutical companies, to vaccine beneficiaries during the nationwide discussion on HPV vaccination between 2009 and 2010 (Figures 2 and 3). After comparing Figures 2 and 3, we defined stakeholders as people or groups who might have some impact or influence on the HPV vaccination situation in Japan. The words describing HPV, features of cervical cancer, vaccine, or the name of the pharmaceutical company appeared with high frequency (Figure 2). The media primarily reported the development or licensure of HPV vaccine in 2009, implying that the main stakeholders were vaccine providers. In contrast, the keywords “public expense,” “subsidy,” “signature campaign,” and an attorney’s name appeared in Figure 3. These keywords were closely related to lay people’s activities or financial issues related to individual vaccination. The increase in the appearance of the keywords implies that the stakeholders in 2010 were lay people who generally benefit from vaccine.

The nature of the discussion regarding HPV vaccination varied widely between newspaper articles and Web pages. We found 2 major tendencies: (1) the number of newspaper articles containing positive and neutral keywords increased exponentially before peaking in June 2010, but the number containing negative keywords remained stable (Figure 4), and (2) the number of Web pages containing all 3 categories of keywords increased gradually over the study period (Figure 5). One possible explanation for this difference is timing. When the Japanese government initiates budgetary compilations, typically in May or June, it allows press clubs—including the major newspaper and television providers, but not online media sources—a high level of access to its proceedings [18]. Thus, the nature of reporting by print and online sources may differ during these periods. Another possible factor is bias. Newspapers are often financially dependent on their sponsors, which may include pharmaceutical companies [19]; thus, they may be reluctant to report information opposing their sponsors’ interests. Another possible factor may be that print media have strict space and time restrictions. Because of these restrictions, newspaper media sometimes report only one aspect of the events.

In contrast, anyone can provide online content, and many of these authors have no conflicts of interest with sponsors. Web content writers, especially amateur writers, can provide their opinions freely on Web. Moreover, because Web content remains almost permanently, they are gradually accumulated.

This discussion does not prove causality, but proposes hypotheses consistent with the observed trends. Another possibility is that, for example, regardless of superficial differences regarding words and causations, the 2 tendencies are interrelated. Their coupling effect can be suggested as a hypothesis, according to the following analysis. The curve in Figure 4 for newspaper articles, in which positive expressions correspond to persuasive messages may superficially look different from the curve in Figure 5, in which negative expressions (ie, criticisms and counter opinions to HPV vaccination) persistently increased on the Web. However, we found a significant covariation between the 2 tendencies. We calculated the values from the equations displayed in Figure 6 to see correlation between the 2 tendencies.

For example, growth of positive news (May 2010) is equal to 30/17, because the value of count positive news (May 2010)—the number of articles with positive words in May 2010, equal to 47 as shown in Figure 3—is larger than 17 by 30, which is the average of count of positive news for the 4 months from January to April 2010. We used 4 months as the denominator in both equations to omit noise that appeared as small variations over periods of 4 months. Then, we extinguished the outlying period January 2010, where growth of positive news (January 2010) was 4.09. This value was exceptionally large in comparison with other periods, where the second largest growth of positive news was 2.4 (SD 1.18). As a result, as in Figure 7, the sequences of growth of positive news (t) and growth of negative Web (t-dt) are correlated significantly for dt=0, where dt means the length of staggered time in evaluating the temporal correlation between the sequences. That is, the strong correlation (R=.79) for dt=0 means the coincidence of changes in the frequencies of words in each category (positive/neutral/negative), between the 2 sequences for exactly same periods (May 2009 to June 2010). Thus, the negative reaction of the Web can be considered to be linked significantly with the positive in the mass media represented by newspapers.

One of the potential explanations for the tendencies is the 2-sided presentation, although its causal relation to the national agreement about HPV vaccination in Japan is difficult to prove until we observe sufficient variation of situations (ie, whether each piece of candidate causal information was shown or not for each change of national agreement—this point is addressed in future work mentioned subsequently). This effect, established by Hovland et al [20], was that a 2-sided message—a message presenting not only the negative side but also the positive side of a controversial issue (such as an advertised product)—is more persuasive than a 1-sided message to educate an audience expected to have basic knowledge about the issue discussed or to those who were not initially positive to be persuaded by a message. This difference between the effects of 1- and 2-sided

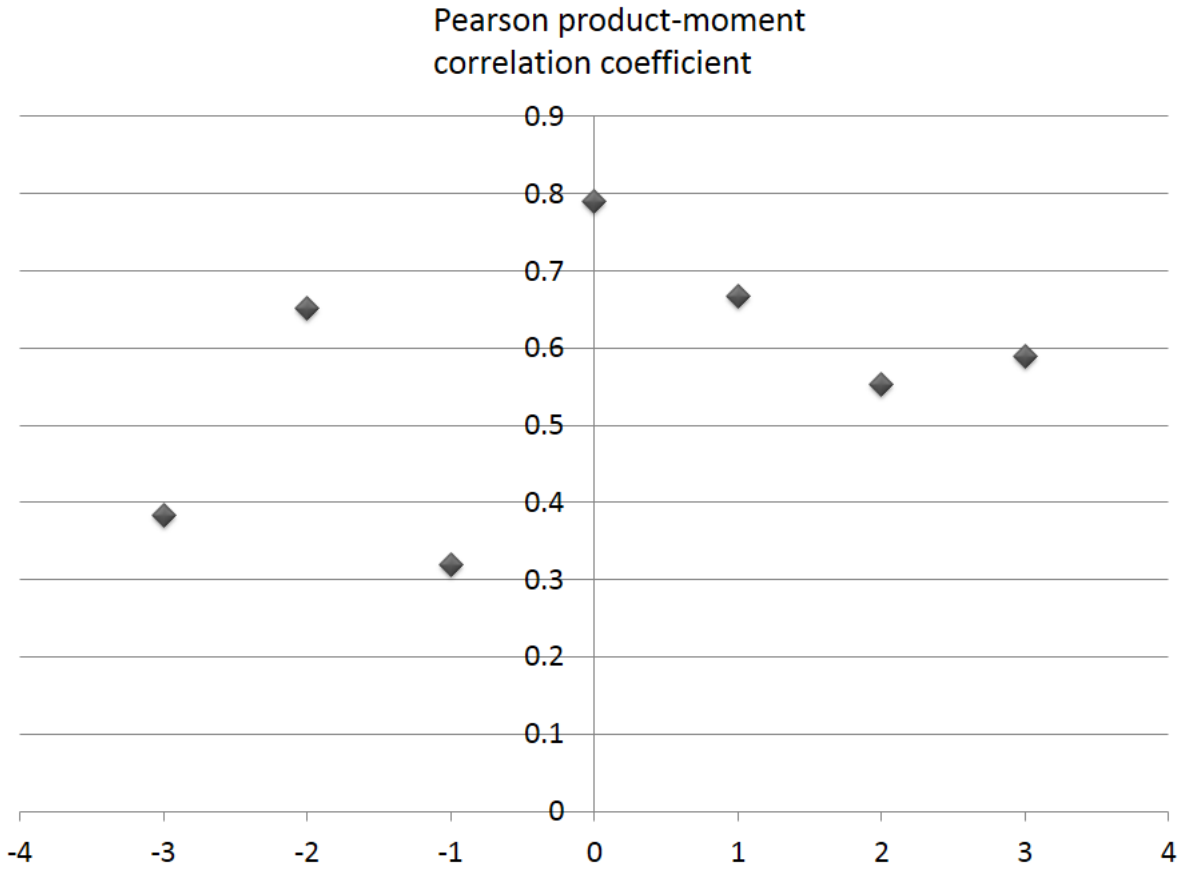
messages is explained in terms of perception of bias in Belch [21]. That is, a 1-sided message would be regarded as biased by those who are aware of opposing arguments or of arguments on both sides of the issue. Taking this position, the analysis result presented here corresponds to this effect in that the positive information provided by newspapers and the negative information on the Web composed 2-sided messages and were casted before and with the social movement toward the promotion of vaccination.

Figure 6. Equations for each month (t) where growth x(t) denotes the increase in the counted number, countx(t) of articles or Web pages corresponding to x at time t, in comparison with the average of the 4 preceding months.

$$growth_{\text{positive\_news}}(t) = \{4\ count_{\text{positive\_news}}(t)/\sum_{i\in[1,4]} count_{\text{positive\_news}}(t\text{-}i)\} - 1 \quad (1)$$

$$growth_{\text{negative\_web}}(t) = \{4\ count_{\text{negative\_web}}(t)/\sum_{i\in[1,4]} count_{\text{negative\_web}}(t\text{-}i)\} - 1 \quad (2)$$

Figure 7. Pearson correlation coefficient between the 2 sequences corresponding to positive keywords in news and negative keywords on the Web, with staggering the sequences by sliding time dt.



Limitations

We have tried to present a plausible explanation for the media trends and national agreement on HPV vaccination. We should consider 2 limitations faced by this study that may limit the generalizability of the findings. First, certain types of media, including magazines, television, and social media (ie, Twitter and Facebook)—the last of which have become increasingly important sources of information transmission between doctors and patients as well as among medical researchers [19,22]—were not included in the analysis because of their lack of inclusion in major databases. Because we do not have a well-established database of social media content, we focused on newspapers and Web pages in this study. Further investigation of social media will be necessary in our next study. Second, pharmaceutical companies might have indirectly influenced the national agreement by promoting HPV vaccination through donations to patient advocacy groups or nonprofit organizations. Because no database was available about advertisements in newspapers and on Web media in Japan, it is difficult to evaluate its influences on the public opinion. Nagata and colleagues [23] reported the effects of advertisements in weekly magazines. According to the study, 6 weekly magazines provided 696 articles and 340 advertisements relating to cancer, 30.4% of which reported dubious folk medicine and immunotherapeutic treatments without supporting evidence. The contents in the weekly magazines could prejudice the public against cancer therapies.

Pharmaceutical companies may intend to influence public opinion via mass media. Both the first and the second limitations are addressed in our future work identifying the role of references to information as factors responsible for the development of national agreement on HPV vaccination. We expect this will present plausible explanations for the positive social trend and support professions in the future who would contribute to prevailing useful medical technologies.

Conclusion

The rapid development of a national agreement regarding HPV vaccination in Japan may be primarily attributed to advocacy by vaccine beneficiaries supported by celebrity advocacy and positive reporting by print and online media.

Acknowledgments

The authors thank Mr Takao Ozawa, Director of Civic Force, and Mr Yohei Kiguchi, CEO of Milog Inc, for their insightful comments. This study is supported by Grant-in-Aid for Scientific Research (No: 23590779).

Conflicts of Interest

None declared.

Multimedia Appendix 1

KeyGraph is visualizing a word map for keyword extraction.

[PDF File (Adobe PDF File), 96KB - jmir_v16i5e129_app1.pdf]

References

1. Paavonen J, Naud P, Salmerón J, Wheeler CM, Chow SN, Apter D, HPV PATRICIA Study Group. Efficacy of human papillomavirus (HPV)-16/18 AS04-adjuvanted vaccine against cervical infection and precancer caused by oncogenic HPV types (PATRICIA): final analysis of a double-blind, randomised study in young women. Lancet 2009 Jul 25;374(9686):301-314. [doi: 10.1016/S0140-6736(09)61248-4] [Medline: 19586656]
2. Nakada H, Murashige N, Matsumura T, Kodama Y, Kami M. Informal network of communication tools played an important role in sharing safety information on H1N1 influenza vaccine. Clin Infect Dis 2010 Oct 1;51(7):873-874 [FREE Full text] [doi: 10.1086/656297] [Medline: 20809846]
3. Lancucki L, Sasieni P, Patnick J, Day TJ, Vessey MP. The impact of Jade Goody's diagnosis and death on the NHS Cervical Screening Programme. J Med Screen 2012 Jun;19(2):89-93 [FREE Full text] [doi: 10.1258/jms.2012.012028] [Medline: 22653575]
4. Murashige N. Japan's immunisation policy in routine, pandemic and post-tsunami situations. Int J Clin Pract 2011 Nov;65(11):1126-1131. [doi: 10.1111/j.1742-1241.2011.02766.x] [Medline: 21995691]
5. The Japan Times Online. 2003. Families win lawsuit over MMR vaccine URL: http://search.japantimes.co.jp/member/member.html?nn20030314a7.htm [accessed 2014-05-08] [WebCite Cache ID 6PPt0Kbl7]
6. Colgrove J, Abiola S, Mello MM. HPV vaccination mandates--lawmaking amid political and scientific controversy. N Engl J Med 2010 Aug 19;363(8):785-791. [doi: 10.1056/NEJMsr1003547] [Medline: 20818883]
7. Larson HJ, Brocard P, Garnett G. The India HPV-vaccine suspension. Lancet 2010 Aug 21;376(9741):572-573. [doi: 10.1016/S0140-6736(10)60881-1] [Medline: 20728739]
8. Ohsawa Y, Benson NE, Yachida M. KeyGraph: Automatic Indexing by Co-occurrence Graph based on Building Construction Metaphor. 1998 Apr 22 Presented at: IEEE International Forum on Research and Technology Advances in Digital Libraries; April 22-24, 1998; Santa Barbara, CA p. 12-18. [doi: 10.1109/ADL.1998.670375]
9. Moubarak G, Guiot A, Benhamou Y, Benhamou A, Hariri S. Facebook activity of residents and fellows and its impact on the doctor-patient relationship. J Med Ethics 2011 Feb;37(2):101-104. [doi: 10.1136/jme.2010.036293] [Medline: 21160080]
10. Ohsawa Y, McBurney P. Chance Discovery. NY: Springer; 2003.
11. Takita M. , Tanaka, Y., Kodama, Y., Murashige, N., Hatanaka, N., Kishi, Y., Matsumura, T., Ohsawa, Y.,Kami, M. Data mining of mental health issues of non-bone marrow donor siblings, Journal of Clinical Bioinformatics 2011;1:19. [doi: 10.1186/2043-9113-1-19]
12. Fruchter R, Ohsawa Y, Matsumura N. Knowledge reuse through chance discovery from an enterprise design-build enterprise data store. New Mathematics and Natural Computation 2005;1(3):393-406.
13. Abe A, Ohsawa Y, Ozaku H, Sagara K, Kuwahara N, Kogure K. Communication error determination system for multi-layered or chained situations. Fundamenta Informaticae 2010;98(1):123-142.
14. Executive Committee for Governmental Subsidies of HPV Vaccine. 2010. [Cervical vaccine grant public funds to women] URL: http://hpv.umin.jp/ [accessed 2014-05-08] [WebCite Cache ID 6PPtxSbL5]
15. Hellosmile. URL: http://www.tfm.co.jp/hellosmile/index.php [accessed 2014-05-08] [WebCite Cache ID 6PPu3EFWf]

16. allwomen.jp. Mirai Happy Project URL: http://allwomen.jp/kitty/index.html [accessed 2014-05-08] [WebCite Cache ID 6PPuEcIvn]
17. Gollust SE, Dempsey AF, Lantz PM, Ubel PA, Fowler EF. Controversy undermines support for state mandates on the human papillomavirus vaccine. Health Aff (Millwood) 2010 Nov;29(11):2041-2046 [FREE Full text] [doi: 10.1377/hlthaff.2010.0174] [Medline: 21041746]
18. Takahara K. The Japan Times Online. 2007. Press clubs: exclusive access to, pipelines for info URL: http://search.japantimes.co.jp/ [accessed 2014-05-08] [WebCite Cache ID 6PPuODTkC]
19. Morton J. American Journalism Review. 1994. Dealing with a conflict of interest URL: http://ajrarchive.org/Article.asp?id=3627 [accessed 2014-05-08] [WebCite Cache ID 6PPuUKyK9]
20. Hovland C, Lumsdaine AA, Sheffield FD. Experiments on Mass Communication. Princeton, NJ: Princeton University Press; 1949.
21. Belch GE. The effects of message modality on one- and two-sided advertising messages. Advances in Consumer Research 1983;10:21-26.
22. Social media and research workflow. London: CIBER, University College London, Emerald Group Publishing Ltd; 2010. URL: http://ciber-research.eu/download/20101111-social-media-report.pdf [accessed 2014-05-08] [WebCite Cache ID 6PPxwVmLE]
23. Nagata M, Takita M, Kishi Y, Kodama Y, Matsumura T, Murashige N, et al. Cancer articles in weekly magazines: useful media to deliver cancer information to the public? Jpn J Clin Oncol 2013 Apr;43(4):426-430 [FREE Full text] [doi: 10.1093/jjco/hyt004] [Medline: 23365111]

Abbreviations

dt: length of staggered time evaluating temporal correlation between sequences
HPV: human papillomavirus

Edited by G Eysenbach; submitted 24.07.13; peer-reviewed by E Yom-Tov, C Harle; comments to author 20.09.13; revised version received 12.12.13; accepted 09.04.14; published 15.05.14.

Please cite as:
Nakada H, Yuji K, Tsubokura M, Ohsawa Y, Kami M
Development of a National Agreement on Human Papillomavirus Vaccination in Japan: An Infodemiology Study
J Med Internet Res 2014;16(5):e129
URL: http://www.jmir.org/2014/5/e129/
doi:10.2196/jmir.2846
PMID:24834471

Original Paper

I Don't Believe It, But I'd Better Do Something About It: Patient Experiences of Online Heart Age Risk Calculators

Carissa Bonner[1,2], MPH; Jesse Jansen[1,2], PhD; Ben R Newell[3], PhD; Les Irwig[1], MBBS, PhD; Paul Glasziou[4], MBBS, PhD; Jenny Doust[4], MBBS, PhD; Haryana Dhillon[2], PhD; Kirsten McCaffery[1,2], PhD

[1]Screening and Test Evaluation Program (STEP), Sydney School of Public Health, The University of Sydney, Sydney, Australia

[2]Centre for Medical Psychology and Evidence-Based Decision-Making (CeMPED), The University of Sydney, Sydney, Australia

[3]School of Psychology, University of New South Wales, Sydney, Australia

[4]Faculty of Health Sciences and Medicine, Bond University, Robina, Australia

Corresponding Author:
Kirsten McCaffery, PhD
Screening and Test Evaluation Program (STEP)
Sydney School of Public Health
The University of Sydney
Rm 301, Edward Ford Building A27
The University of Sydney
Sydney, NSW 2006
Australia
Phone: 61 2 9351 7220
Fax: 61 2 9351 5049
Email: kirsten.mccaffery@sydney.edu.au

Abstract

Background: Health risk calculators are widely available on the Internet, including cardiovascular disease (CVD) risk calculators that estimate the probability of a heart attack, stroke, or death over a 5- or 10-year period. Some calculators convert this probability to “heart age”, where a heart age older than current age indicates modifiable risk factors. These calculators may impact patient decision making about CVD risk management with or without clinician involvement, but little is known about how patients use them. Previous studies have not investigated patient understanding of heart age compared to 5-year percentage risk, or the best way to present heart age.

Objective: This study aimed to investigate patient experiences and understanding of online heart age calculators that use different verbal, numerical, and graphical formats, based on 5- and 10-year Framingham risk equations used in clinical practice guidelines around the world.

Methods: General practitioners in New South Wales, Australia, recruited 26 patients with CVD/lifestyle risk factors who were not taking cholesterol or blood pressure-lowering medication in 2012. Participants were asked to “think aloud” while using two heart age calculators in random order, with semi-structured interviews before and after. Transcribed audio recordings were coded and a framework analysis method was used.

Results: Risk factor questions were often misinterpreted, reducing the accuracy of the calculators. Participants perceived older heart age as confronting and younger heart age as positive but unrealistic. Unexpected or contradictory results (eg, low percentage risk but older heart age) led participants to question the credibility of the calculators. Reasons to discredit the results included the absence of relevant lifestyle questions and impact of corporate sponsorship. However, the calculators prompted participants to consider lifestyle changes irrespective of whether they received younger, same, or older heart age results.

Conclusions: Online heart age calculators can be misunderstood and disregarded if they produce unexpected or contradictory results, but they may still motivate lifestyle changes. Future research should investigate both the benefits and harms of communicating risk in this way, and how to increase the reliability and credibility of online health risk calculators.

(J Med Internet Res 2014;16(5):e120) doi:10.2196/jmir.3190

KEYWORDS

cardiovascular disease; prevention; risk calculator; risk assessment; risk perception; lifestyle; behavior change

Introduction

Health risk calculators are widely available on the Internet, with outcomes ranging from overall mortality to specific diseases such as cancer, diabetes, and cardiovascular disease (CVD) [1]. In the context of CVD, risk calculators use a mix of clinical and lifestyle risk factors to estimate the probability of a heart attack, stroke, or death over a specific period of time [2]. This may be communicated as a percentage or the alternative risk format of "heart age", where heart age older than current age indicates modifiable risk factors [3,4]. These calculators may impact patient decision making about CVD risk management with or without clinician involvement, but little is known about how patients use and understand such risk calculators. Previous research on diabetes and cancer risk calculators suggests that people may disregard results that do not match their prior risk perception [1,5], and the presented numerical format may affect perceived credibility of the results [6].

Clinical guidelines around the world advocate CVD risk assessment based on "absolute risk"—the percentage risk of a cardiovascular event over a 5- or 10-year timeframe [7]. The Framingham model is commonly used and accounts for the effect of non-modifiable risk factors, including age and gender, as well as modifiable risk factors, such as smoking, blood pressure, and cholesterol [8,9]. However, research has established that percentages are poorly understood by both clinicians and patients [10,11]. Clinicians also report situations in which communicating absolute risk to patients is unhelpful [12,13]. In particular, patients with lifestyle risk factors (eg, smoking or obesity) can have low percentage risk (eg, younger patients and women are likely to have low 5-year absolute risk<10%), which may reduce motivation to change lifestyle before it leads to CVD and other chronic illnesses [12]. Such communication issues may discourage GPs from using absolute risk assessment, contributing to the suboptimal use of absolute risk guidelines around the world [14,15].

Preliminary research suggests that converting percentage risk into an individual's heart age may be a useful alternative for communicating CVD risk. A focus group study using hypothetical risk found that patients preferred heart age over other CVD risk formats, but there were concerns it may frighten people if older than their current age [16]. A randomized controlled trial found that giving patients a CVD risk profile, including heart age, improved cholesterol levels compared to usual care over the first year of cholesterol medication treatment, especially for higher risk patients [17]. The similar concept of "lung age" was found to motivate smokers to quit regardless of the result: normal lung age acted as an incentive to stop smoking and abnormal lung age sent a message that quitting could slow deterioration [18]. However, heart age and lung age were not specifically compared to percentage risk in these studies [17,18]. An experimental study found that heart age improved understanding of risk compared to 10-year percentage risk and had more emotional impact for younger people at higher risk [19]. A study on a New Zealand heart age tool suggests it may increase clinician understanding and confidence in assessing absolute CVD risk, but patient outcomes were not assessed [3].

To the authors' knowledge, there have been no studies investigating patient understanding of heart age compared to 5-year percentage risk, which is currently used in Australian guidelines and online tools. Nor have there been any studies investigating the best way to present heart age. This study aimed to investigate patient experiences and understanding of online heart age calculators that use different verbal, numerical, and graphical formats, based on 5- and 10-year Framingham risk equations [8,9], which are used in clinical practice guidelines in many countries around the world [3,7,20,21].

Methods

Ethical Approval

Ethical approval for the study was obtained through the Human Research Ethics Committee of the Sydney Local Health District (Protocol No. X11-0200). Each participant gave written consent before participating in the interview.

Recruitment

General practitioners (GPs) in New South Wales, Australia, recruited 26 patients between 40-70 years of age, with at least one CVD or lifestyle risk factor, who were not currently taking medication, targeting low (5-year absolute risk<10%) to moderate (10-15%) risk patients who may be less motivated by their percentage risk result [20,21]. Purposive sampling was used to recruit participants with a range of ages, gender, knowledge of risk factors, and risk calculator results (see Table 1). This was done by modifying the eligibility criteria given to recruiting GPs throughout the study. Analyses based on 26 participants suggested saturation of key themes (see Figure 1), so no further recruitment was conducted [22].

Table 1. Participant characteristics in order of absolute risk result, by gender.

ID	AR[a]	Gender	Age	HA[b]: Unilever[c]	HA: New Zealand[d]	HA vs age	Knew SBP[e]	Knew Chol[f]
102	1%	Woman	48	46	<48	younger	Y	Y
109	1%	Woman	52	46	<52	younger	Y	N
112	1%	Woman	54	49	<54	younger	Y	N
99	2%	Woman	40	51	64	older	N	N
118	2%	Woman	51	51	61	mixed	N	N
68	3%	Woman	57	43	59	mixed	Y	Y
87	4%	Woman	57	61	64	older	Y	Y
115	4%	Woman	63	62	69	mixed	Y	Y
108	4%	Woman	67	62	68	mixed	Y	N
103	6%	Woman	39	39	39	same	N	N
71	6%	Woman	57	53	<57	younger	Y	Y
70	6%	Woman	58	74	72	older	Y	Y
107	8%	Woman	49	59	60	older	Y	Y
119	8%	Woman	59	79	70	older	N	N
116	9%	Woman	60	80	73	older	N	Y
106	10%	Woman	66	72	>75	older	N	N
84	3%	Man	45	51	52	older	Y	N
91	3%	Man	48	52	58	older	N	N
111	4%	Man	50	60	57	older	Y	N
63	5%	Man	55	57	66	older	N	N
96	5%	Man	58	63	62	older	Y	Y
113	6%	Man	41	43	46	older	Y	N
110	7%	Man	62	58	63	mixed	Y	Y
94	8%	Man	60	74	69	older	Y	N
95	11%	Man	58	60	70	older	N	N
65	12%	Man	55	60	66	older	N	Y

[a]AR: initial 5-year absolute risk estimate on New Zealand calculator (<10% indicates low risk; 10-15% indicates moderate risk)

[b]HA: heart age result on each website

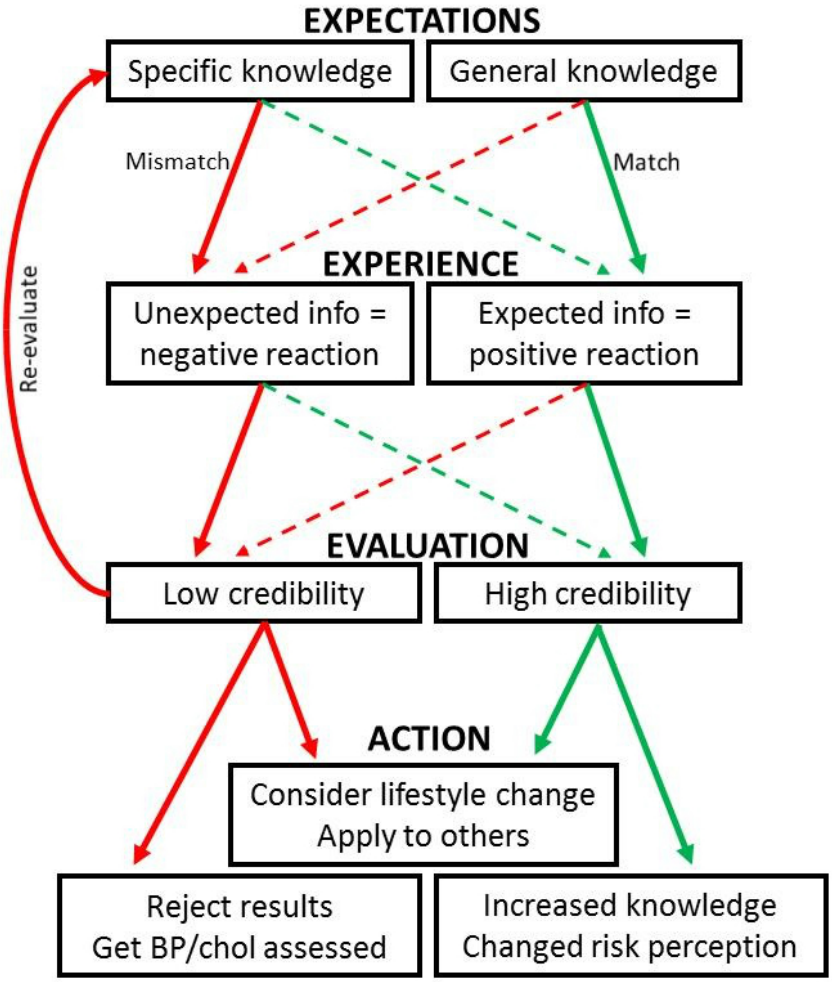
[c]Unilever: website developed by Unilever [23]

[d]New Zealand: website developed by New Zealand Heart Foundation [24]

[e]SBP: systolic blood pressure, Y=yes, N=no

[f]Chol: cholesterol, Y=yes, N=no

Figure 1. Process of using risk calculators: red arrows indicate low credibility pathways, green arrows indicate high credibility pathways, solid lines indicate main pathways identified, dashed lines indicate alternative pathways identified.



Participants

Participants were 16 women and 10 men, between 40 and 67 years of age, with a range of highest educational attainment: 4 had not completed high school, 6 had completed high school, 7 had a technical qualification, and 9 had completed a university degree. Five-year absolute risk results ranged from 1-12%, with 23 at low risk (<10%) and 3 at moderate risk (10-15%) of a CVD event. Compared to current age, the heart age results were: 16 older, 4 younger, 1 same as current age, and 5 mixed results for the two calculators.

Process

Two heart age calculators based on Framingham risk equations were publicly available at the time of the study in 2012 (see Table 2 and Figure 2) [23,24]. A protocol including think aloud and semi-structured interview methods was developed based on past research showing that a concurrent think aloud protocol elicits more information, but additional insights can be gained retrospectively (see Multimedia Appendix 1 for protocol) [25]. The interviewer (CB) was a researcher trained in public health qualitative methods, who piloted the protocol with a convenience sample of 4 participants who met the study eligibility criteria. Pilot participant feedback was discussed with the research team and the think aloud instructions were clarified before commencing the study. Participants were asked to think aloud as they used each website in random order, with minimal input from the interviewer unless they had difficulty using the website. No interpretation of the results was provided until the end of the interview and the interviewer clarified that the researchers were not connected to the websites if this issue arose. In order to practice thinking aloud consistently, participants completed a simple “spot the difference” task in which they described what they were doing. Upon successful completion of this practice task, participants began using the heart age calculators. A “keep talking” sign was placed above the computer, which the interviewer would point to if the participant was silent for more than 10 seconds. The entire session was audio-recorded and use of the websites was video-recorded using SMRecorder screen capture software [26]. All interviews were audio-recorded, but technical problems prevented the use of screen capture software in some interviews: 3 participants had no video data because of computer or software issues and 5 participants had only one calculator video-recorded (3 due to computer or software issues and 2 due to the Unilever website being taken down, who completed the study over the phone when the website became temporarily available again).

Figure 2. Example of heart age calculator results for ID91: male, age 48, BP and cholesterol unknown. A: New Zealand, initial absolute risk result 3% but estimate increased to 5%; B: heart age 58; C: Unilever result, heart age 52.

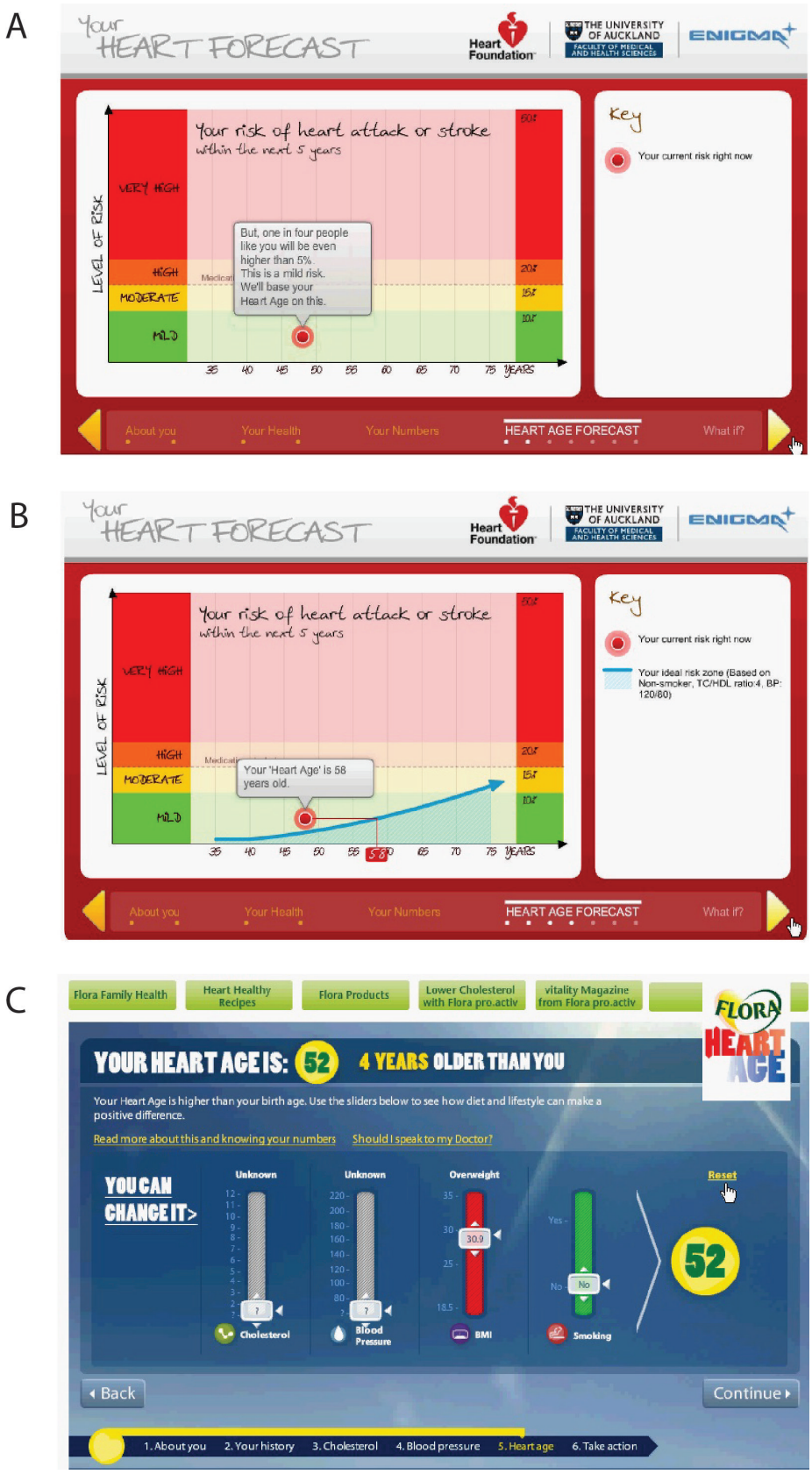


Table 2. Main differences between the two heart age calculators.

Variable	New Zealand Heart Foundation [24]	Unilever [23]
Timeframe	5-year risk based on cholesterol Framingham risk equation [9].	10-year risk based on cholesterol or body mass index Framingham risk equation [8].
Minimum heart age	"Lower than" current age. Current age is the lowest value shown.	Heart age result can be younger than current age.
Include % risk	Yes – % and risk level (mild, moderate, high, very high).	No – heart age only.
Graphical display	Yes – trajectory over age with colors indicating risk level.	No – text only.
Risk factors asked about	Age, gender, family history, smoking, diabetes, systolic blood pressure, cholesterol ratio, ethnicity.	Age, gender, family history, smoking, diabetes, systolic/diastolic blood pressure, total/HDL cholesterol, height, weight, waist.
Modifiable risk factors at final results page	Blood pressure, cholesterol, smoking, diabetes (if not already diagnosed).	Blood pressure, cholesterol, smoking, body mass index.
Missing data	If blood pressure and/or cholesterol values were not known, two values were given: a population average based on demographics, and a higher than average value that "1 in 4 people like you" would have. These estimates were used to calculate two absolute risk results (see Table 1 for initial result based on the average). The higher than average value was used to calculate heart age.	If blood pressure and/or cholesterol values were not known, alternative Framingham-based algorithms including body mass index were used to calculate heart age [8].

Analysis

A framework analysis method was used to analyze the interview transcripts, which involved five steps [27]. The first step was familiarization with the data: CB read through all 26 transcripts, recorded the calculator input/output for each participant from videos if available or from transcripts and field notes if unavailable, and discussed this and 2 transcript excerpts with all authors, covering younger and older heart age results. The second step was to create a thematic framework: CB, JJ, and KM read a sample of 5 transcripts covering younger, same/mixed, and older heart age results, and developed the initial framework. Third was indexing: CB, JJ, and HD each watched a video to ensure understanding of the process and coded the remaining transcripts according to the framework, with new themes and revisions to the framework discussed (see Multimedia Appendix 2 for final framework). Ten transcripts were double-coded independently by two researchers. The fourth step in the analysis was charting: CB, JJ, and HD summarized themes and supporting quotes from each transcript in the framework (a matrix with participants as rows and themes as columns), with transcripts reread and discussed to resolve any disagreement about the best way to represent the data. The fifth step was mapping and interpretation: CB and JJ examined the framework within and across themes and participants to identify overarching themes and relationships, independently summarized the process of using and interpreting the two risk calculators in a diagram, and differences between the 2 process diagrams were discussed with KM and HD. The order in which the calculators were viewed was taken into account, based on separate coding of the two websites, but this did not appear to influence the overall process. Then JJ, HD, and KM each read 2 additional transcripts to check the findings and the final process diagram was created (see Figure 1). The final results were discussed with all authors. Rigor was addressed by: repeated coding of transcripts by different team members to ensure a comprehensive themes list and framework was achieved; an iterative process of constant comparison between the existing framework and new data; detailed documentation of the analysis process; and discussion of emerging and final themes with all authors [28].

Results

Reliability

The reliability of the risk calculator results was affected by several factors. Misinterpretation of risk factor questions was common, with 9/26 participants making at least one error in their responses to the questions, based on inconsistencies between the data entered in the two calculators and their thought process while entering the data. Many participants couldn't remember their exact blood pressure or cholesterol levels but were aware that these risk factors were low or normal based on past assessment, and this did not match the estimates provided by the calculator.

> *I guess if I'd known my blood pressure and cholesterol level, it might have been lower but that's all right…automatically went to the default position, the higher setting… I think I'd be lower than that in reality.* [ID95, older heart age]

The two calculators gave different heart age results for all but one participant due to these input errors and the different assumptions built into the calculators: the use of different absolute risk models (5- vs 10-year risk; use of body mass index if cholesterol was unknown), the estimates used when blood pressure was unknown, and the "ideal" that participants were compared to when calculating the heart age.

Risk Formats

In the New Zealand calculator, the explanation of the percentage risk information was often overlooked, with 12/26 participants skipping through at least one part of the explanation. Even when the percentage information was seen, it was often forgotten by

the end, with participants generally more focused on the graph and heart age results. When specifically asked, many participants were unsure what the percentage referred to, even though the graph title clearly stated it was the percentage risk of a heart attack or stroke in the next 5 years.

> *Well, that I have a 2%, so I have 2. Well, what does it mean? Does it mean that I, 2 days out of 100, I'm at risk of a heart attack. I don't know what that means. I have a 2% chance, I have a…well, it sounds low but what does it mean? I mean I don't know…No, I think the heart age was good.* [ID118, mixed heart age results]

The graph in the New Zealand calculator was sometimes confusing, particularly the first few screens where many different numbers were used in the explanation of the results (see Figure 2A). However, other participants liked the graph and demonstrated a clear understanding of how risk would increase with age when viewing it.

> *Yeah, I think the graph is not so great…it is more the older generation I guess you'd be looking at as well, not people who are really computer savvy and they can read 100 things on the screen at one time and take it all in.* [ID107, older heart age]

> *The risk is going to go up anyway and even if I stay doing what I'm doing now, it's still going to rise, and to keep it at the lowest risk I need to do more…being vigilant on the blood pressure and the cholesterol and just being probably a bit more active.* [ID115, mixed heart age results]

The heart age format in both calculators was much more salient, meaningful, and elicited more emotional reactions in participants. Younger heart age was seen as very positive and participants often preferred the calculator that gave them a lower heart age, although some felt it was unrealistic and could discourage people from improving their lifestyle. Older heart age was confronting and participants' first reaction was often disbelief, particularly if they felt they had a good lifestyle. Heart age had a more immediate meaning to participants than percentage risk, indicating a healthy lifestyle if younger, and the need to change if older.

> *Oh, younger than me—that's good news (laughter)…The fact that it's younger. I mean I already feel that I am probably healthy-ish for my age, healthy. So I'm not assuming that yes, ok, my heart is 6 years younger than my body but I'm, to me that says, yeah, you're ok.* [ID109, younger heart age]

> *Your heart age is 52? Get out of town…How does that work? Your current risk right now is there. But my heart age is 52…No cholesterol. Normotensive…Don't believe it…not smoking, eating a healthy diet…I'm grumpy with this website already. Because it's asking me to do things that it didn't actually question me about before, like being active or eating.* [ID84, older heart age]

The multiple formats presented in the New Zealand calculator were sometimes perceived as contradictory, particularly low percentage risk in the green mild category compared to an older heart age result. The use of consistent colors was also important—green indicated a mild risk level in the basic graph, but a green line was also used to show how risk would increase if the participant started smoking or developed diabetes.

> *Well, off the basis of one number, you're saying I'm middle in the mild, yet you've automatically put me at 5 years older than my heart. So those two things are probably contradictory in a way…by clicking on the start smoking you strangely get a green line, which would indicate a good thing which is probably not right. Ah…it should be a black line.* [ID113, older heart age]

The ability to modify different risk factors to see what effect they had on the results also had mixed responses from participants: some weren't interested in using it or didn't understand it, while others spent some time playing around with the factors to reduce their heart age result or make the estimated risk factors more realistic. However, the message this conveyed depended on how the participant used it—for example, the following participant concluded that blood pressure was more important than cholesterol because she happened to move blood pressure to a higher level.

> *What about if you get diabetes and you smoke? You are dead by 55. Wow…Oh, so the cholesterol isn't too bad. It's, gauging from this, it's when the blood pressure goes up and if you have diabetes. That's what I'm getting from this.* [ID103, same heart age as current age]

Participants tended to focus on the risk factors that were most relevant to them; for example, the effect of smoking was of more interest to current/ex-smokers and several participants wanted to see the effect of alcohol.

> *Your projected risk if no changes are made…ok, so I go from a mild to a high if I don't change anything that I do. All right…If I quit smoking long term…that decreases it…so I'm only in the moderate range then if I do that, ok…It's something that I have actually been thinking of for quite a while.* [ID99, older heart age]

> *What about drinking, where is the drinking? That would be more interesting to me…you would have to put drinking in there as we get older.* [ID70, older heart age]

Process of Using Risk Calculators

The process of using the risk calculators involved several stages: expectations of CVD risk based on prior knowledge, experience of using the calculator, evaluation of the credibility of the results, and actions considered as a result of this process. This is illustrated in Figure 1, with examples in Tables 3 and 4. Participants' expectations, experience, evaluation, and actions sometimes changed between the two calculators, but the order in which they were viewed did not appear to affect the overall process. The process diagram therefore describes the range of pathways that participants followed regardless of the order in which the calculators were used, with solid lines indicating the

two main pathways, and dashed lines indicating alternative pathways.

The high credibility pathway tended to occur when participants had little prior knowledge of their own risk of CVD. Their general knowledge of CVD risk factors was consistent with the information in the calculators, and so they tended to accept the results and have a more positive reaction to them. Those who received a younger heart age result usually had a positive reaction. In these situations the credibility of the results was not closely questioned and led to various actions: participants considered changing their lifestyle to lower risk even further or maintain younger heart age, thought of higher risk family or friends who could benefit from using the calculator, and sometimes perceived increased understanding or changed risk perception.

Alternatively, those with more specific knowledge evaluated the calculator as having high credibility if they had a positive reaction to the calculator, such as getting lower heart age than current age or all the information matched what they knew about their CVD risk. Seeing a similar result for the second calculator also increased credibility.

The low credibility pathway tended to occur when participants had more specific knowledge of their own risk of CVD, but could not necessarily remember their exact blood pressure and cholesterol levels. These participants were more likely to encounter unexpected information in the calculators, and reacted negatively to receiving an unexpected heart age result. In these situations, the credibility of the results was questioned and participants re-evaluated their prior expectations (eg, that they already had a good diet) and experience of using the calculator (eg, that they weren't asked any questions about diet) to make sense of the result. Common reasons used to discredit the results were the lack of relevant lifestyle questions in the New Zealand calculator and the influence of corporate sponsorship in the Unilever calculator. However, even when the results were rejected, participants still considered lifestyle change and felt that the calculators would be useful for others. Some decided they should get their blood pressure or cholesterol checked again to increase the accuracy of the risk assessment.

Alternatively, those with little prior knowledge evaluated the calculator as having low credibility if they had a very negative reaction to the results or specific components (eg, some men believed that body mass index was inaccurate for them). However, they still considered lifestyle change and getting a more accurate assessment.

In summary, using the heart age calculators appeared to lead participants to consider lifestyle changes regardless of the pathway they described and regardless of their heart age result. This is illustrated by the quotes in Tables 3 and 4.

Table 3. Examples of main pathways for low vs high credibility.

Main pathway for low credibility	Example: quotes from ID70 (woman aged 58, higher heart age, 5-year absolute risk 6%)	Main pathway for high credibility	Example: quotes from ID119 (woman aged 59, higher heart age, 5-year absolute risk 8%)
Expectations: Specific knowledge of own risk factors is less likely to match experience of using calculator	*Do you know your blood pressure – yes. Oh, well…I can't really remember but I'll just put in I think it was 138 over 81…Do you know your cholesterol – can't remember. Oh, hang on a sec, 3 to 4…high, I think it was high…She said it was sort of middle – 5.*	Expectations: Having only general knowledge about CVD is more likely to match experience of using calculator	*I like to think that I am low (risk) but I don't know. My mum did have some issues with her heart when, when she was young, I mean probably, oh, late 60s, early 70s. So, if it is something to do with genetics or whatever, well I'm getting into that age, so I don't know. I would say I'm in low. I would like to believe that.*
Experience: Negative emotional reaction when calculator doesn't match expectations, leading participants to question credibility	*Current risk this is a mild – yeah…I agree with that...Your heart age is what, rubbish. I don't believe that…no way…What…I better start changing this, hadn't I…Too in your face, I don't want to know that. I don't want to know my heart age…72 years of age. Too frank… Terrible, 72...It sounds like I'm going to have a heart attack very soon...I'm on the way.*	Experience: If information makes sense and matches expectations, and/or elicits a positive reaction, credibility is not questioned	*You will be near this point, this is a mild risk, oh good (laughter)….I'm glad about that, happy, happy, happy…this risk will be your risk as you get older, ok. So I have to be careful what I do…Your ideal risk based on non-smoker, your heart age is 70 years old, oh wow (laughter)… You can reduce your risk of heart attack or a stroke by not smoking – I don't, eating a healthy diet – ok, by being active for at least 30 mins on most, on most days of the week – ok, I need to do that.*
Evaluation: Low credibility leads participants to re-evaluate expectations and experience	*Terrible, again it's terrible. I hate this 74 and 72, that's not real…The only one that can say my heart age would be the, my cardiologist when he goes in and has a look at my heart age…A lot of things on the Internet really aren't sort of factual...I reckon other people my age would probably be on a higher…two of my school chums would be definitely because they're overweight, they're on the tablets.*	Evaluation: High credibility leads participants to consider several possible actions without re-evaluating expectations and experience	*Wow, this is very good…it's an eye opener to, you know, I think that I didn't have, to be honest, I didn't have much problem with my heart…Oh yeah, I'm overweight and this and that but never thinking that it's, it (would) have such an effect on my heart and that, yeah, I'm like anybody else. You know I can have a heart attack and I can have problems with my heart, scary…So the cholesterol level and the blood pressure is something that…I need to be very much aware and try to you know make sure that I check it all the time with the Doctor.*
Action: Lifestyle change and usefulness for others considered	*Yes, I think it is a wakeup call. Yes, I am watching my diet. Yes, I am not exercising enough yet but I will (laughter)… I think I might put Mum on here and give her a go.*	Action: Lifestyle change and usefulness for others considered	*My husband should do this…it's telling me that I need to do something. That I have to take action…I don't know if, if the heart can get back to, to match my age...that is something I need to talk to my Doctor about.*

Table 4. Examples of participants considering lifestyle change by heart age result.

Heart age result	Quote illustrating consideration of lifestyle change
Older heart age than current age	*Ok, so this is interesting...if I reduce my BMI...so weight is a clear factor here…This is quite good, this tool here, because it actually gives me some targets for my BMI and what sort of weight I should be. So it's, man, I'm going to have to lose a lot of weight (laughter)...This is quite good because I think it, it clearly shows that my weight is something I need to work on. And I think the fact it's red, it takes you straight to that and I do like that.* [ID65, older heart age]
Mixed results (younger then older heart age result)	*With that graph specifically, that it's a general rise anyway without taking into account, you know it's not going to be stable. It's automatically going up so you have to work a bit harder at it…that's made an impression on me and that's the biggest thing I've picked up that, yeah, you've still got to keep working at it. It doesn't matter what you're doing now that sort of just yeah maintain or get more, just to try and yeah reduce the risk, keep going to reduce it…just being aware of it and, and I think it, I'm pleased that it's at the lower level but also you've got to be vigilant.* [ID115, mixed results]
Same heart age as current age	*When I went on the Heart Foundation one and I changed my cholesterol and it increased my risk of heart disease...that's something that's important to me, because it happens in my family, so, yeah. This sort of thing that I have sitting out here (biscuits) will not be happening. Well, it still will, but not to the extent that it does in this household...I would probably, you know, take out maybe one load of biscuits and put some carrots in.* [ID103, same heart age as current age]
Younger heart age than current age	*So if I move the cholesterol down to 4 and…that reduces me down to 48. So I think I better get myself cracking and...get my cholesterol down... I think it means I'm probably tracking ok, I'd prefer if it was lower so then you know that my cholesterol can get reduced. So I know how to do that it's just that I haven't done it...I like the idea of it being, I like the idea of it being 48 better than 53.* [ID71 , younger heart age]

Discussion

Findings and Implications

Our findings suggest that online heart age calculators prompt people to consider improving their lifestyle regardless of the accuracy and perceived credibility of the results, or the result they received. This is consistent with the findings of the "lung age" study, where both positive and negative results prompted smokers to consider quitting [18]. As found in previous studies on diabetes and cancer risk calculators, participants often disregarded unexpected or negative information [1,5], and actively sought reasons to discredit the results. However, participants in this study considered lifestyle change and felt the calculators would be useful for other people, even when they claimed to disbelieve the calculator result. The value of such tools is therefore dependent on their goal—if it is to prompt people to think about improving their lifestyle regardless of the level of risk [4], they appear to be achieving this; if it is to provide accurate information and understanding of risk, then they could be improved [1]. However, thinking about lifestyle changes does not necessarily translate into actual behavior change. Previous research has found that personalized risk calculators have limited impact on behavior [29], but they may prompt people to seek further information and support as an interim step before behavior change [30].

The ultimate goal of risk calculator websites will vary depending on the motivation and target audience of the organization that created it. Our findings show wide variation in the way that people use and understand features of personalized risk calculators, with implications for designers of such tools. Patients in our study misinterpreted risk factor questions, entered data inconsistently between different risk calculators, skipped information that may have prevented these issues, and did not use all features of the tools. Some of these issues may be improved by simple design alterations, such as larger font, easier navigation, and clearer instructions for how to use the risk factor modification tools. Alternatively, simpler, less interactive information formats may be more effective in terms of information processing and understanding of risk [31,32].

This study is consistent with previous findings that heart age elicits more emotional reactions and is more meaningful to patients than percentages [16,19]. All participants demonstrated good understanding of the effect of risk factors on heart age and its link to lifestyle, but the percentage information was interpreted in many different ways and often overlooked. This supports the large body of research showing that percentages are often misunderstood [10,11]. The downside of heart age was that it was very confronting for participants to receive a much older heart age than their current age and some explicitly said they would prefer not to know.

The presentation of multiple formats was problematic because low percentage risk and older heart age were perceived as contradictory, suggesting that the focus should be on explaining one or the other, not both. Although the New Zealand calculator was developed with a step-wise structure that attempts to fully explain how the results were calculated [33], the large amount of information with multiple numbers was confusing for many participants, and the simplified Unilever format was often preferred even though corporate sponsorship reduced its credibility. The graph used in the New Zealand calculator with color-coded risk categories and projected risk over age (see Figure 2) appeared to be useful additional information for many but not all participants. Since preferences for and understanding of different CVD risk formats were variable in our study, quantitative research is needed to test the effect of presenting heart age in different formats and identify the best way to present such information to different groups of people. Future research should include measurement of benefits like understanding and motivation to improve lifestyle, but also the potential harms of conveying such emotive information, including worry and seeking unnecessary tests for low risk.

Future research on online health risk calculators could also investigate how to increase their reliability and credibility. One option is to involve clinicians in explaining the results, which could improve perceived credibility and the accuracy of the risk factor data, and may prevent misunderstandings. However, since health consumers use online risk calculators outside of clinical consultations [1], it would be beneficial to improve the format of online health information so that unexpected or absent information is fully explained. This should include an explanation of why different risk calculators may produce different results, to increase awareness of accurate data entry for risk factors, and understanding of the assumptions behind the calculation [2]. Our findings suggest that people expect to be asked about broader aspects of their lifestyle and history than those included in the CVD risk models; consequently, the face validity is reduced by exclusion of these questions. The link between lifestyle advice and the risk factor questions also needs to be clear to avoid a negative reaction to the calculator. However, the aim of risk calculator tools will vary depending on the goals of the organization that develops them and the audience they are targeting, so our findings may have different implications for different designers.

Strengths and Limitations

The strengths of this study include a novel topic and rigorous qualitative methods, including purposive sampling to achieve a heterogeneous patient sample, theme saturation, a trained interviewer, use of both semi-structured interview and think aloud data, and a framework analysis process that used multiple analysts and an iterative process to arrive at final themes. The external validity of the study is strengthened by the use of existing online risk calculators, self-reported risk factors, and widely used, validated Framingham risk equations. The limitations include missing video data for some participants and the possibility that the interview questions and presence of the interviewer may have influenced reactions to the websites. However, audio recordings were obtained for all participants, most users of such websites would have a prior interest in CVD, and the interviewer took care to avoid giving any interpretation of the results or reactions to the websites. The results may not reflect how consumers use risk calculators in a more realistic setting, and as typical with qualitative research, the sample was not designed to be representative of the general population but rather present a range of perspectives.

Conclusions

Our findings demonstrate an interesting paradox: online heart age calculators are easily misunderstood and the results may be dismissed if the information is unexpected or negative, but the process of using such calculators may motivate lifestyle change regardless of the outcome. Future research should investigate both the benefits and harms of communicating risk in this way and how to increase the reliability and credibility of online health risk calculators.

Acknowledgments

The authors would like to thank the patients who participated in the study, the GPs and Shannon McKinn for assistance with recruitment, and Mark Cobain for extending public access to the Unilever website after it was taken down to enable data collection for the last two participants.

The study was funded by the National Health and Medical Research Council (NHMRC), through grants awarded to the Screening and Test Evaluation Program (No. 633003) and Bond University (No. 511217). The funders had no role in the design or conduct of the study; in the collection, analysis, and interpretation of the data; or in the preparation or approval of the manuscript.

Authors' Contributions

CB contributed to study design, recruitment, data collection, framework analysis, interpretation, drafting, and revising the manuscript. JJ contributed to study design, framework analysis, interpretation, and revising the manuscript. BN contributed to study design, interpretation, and revising the manuscript. LI contributed to study design, interpretation, and revising the manuscript. PG contributed to study design, interpretation, and revising the manuscript. JD contributed to study design, interpretation, and revising the manuscript. HD contributed to framework analysis, interpretation, and revising the manuscript. KM contributed to study design, framework analysis, interpretation, and revising the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Protocol for data collection: think aloud method and semi-structured interview schedule.

[PDF File (Adobe PDF File), 427KB - jmir_v16i5e120_app1.pdf]

Multimedia Appendix 2

Structure of final analysis framework.

[PDF File (Adobe PDF File), 203KB - jmir_v16i5e120_app2.pdf]

References

1. Harle CA, Downs JS, Padman R. A clustering approach to segmenting users of Internet-based risk calculators. Methods Inf Med 2011;50(3):244-252. [doi: 10.3414/ME09-01-0080] [Medline: 20300681]
2. Allan GM, Nouri F, Korownyk C, Kolber MR, Vandermeer B, McCormack J. Agreement among cardiovascular disease risk calculators. Circulation 2013 May 14;127(19):1948-1956. [doi: 10.1161/CIRCULATIONAHA.112.000412] [Medline: 23575355]
3. Wells S, Kerr A, Broadbent E, MacKenzie C, Cole K, McLachlan A. Does your Heart Forecast help practitioner understanding and confidence with cardiovascular disease risk communication? J Prim Health Care 2011 Mar;3(1):4-9. [Medline: 21359254]
4. Cooney MT, Vartiainen E, Laatikainen T, De Bacquer D, McGorrian C, Dudina A, SCORE and FINRISK investigators. Cardiovascular risk age: concepts and practicalities. Heart 2012 Jun;98(12):941-946. [doi: 10.1136/heartjnl-2011-301478] [Medline: 22626902]
5. Scherer LD, Ubel PA, McClure J, Greene SM, Alford SH, Holtzman L, et al. Belief in numbers: When and why women disbelieve tailored breast cancer risk statistics. Patient Educ Couns 2013 Aug;92(2):253-259. [doi: 10.1016/j.pec.2013.03.016] [Medline: 23623330]
6. Witteman HO, Zikmund-Fisher BJ, Waters EA, Gavaruzzi T, Fagerlin A. Risk estimates from an online risk calculator are more believable and recalled better when expressed as integers. J Med Internet Res 2011;13(3):e54 [FREE Full text] [doi: 10.2196/jmir.1656] [Medline: 21908265]
7. Ferket BS, Colkesen EB, Visser JJ, Spronk S, Kraaijenhagen RA, Steyerberg EW, et al. Systematic review of guidelines on cardiovascular risk assessment: Which recommendations should clinicians follow for a cardiovascular health check? Arch Intern Med 2010 Jan 11;170(1):27-40. [doi: 10.1001/archinternmed.2009.434] [Medline: 20065196]
8. D'Agostino RB, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. Circulation 2008 Feb 12;117(6):743-753 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.107.699579] [Medline: 18212285]
9. Anderson KM, Odell PM, Wilson PW, Kannel WB. Cardiovascular disease risk profiles. Am Heart J 1991 Jan;121(1 Pt 2):293-298. [Medline: 1985385]
10. Akl EA, Oxman AD, Herrin J, Vist GE, Terrenato I, Sperati F, et al. Using alternative statistical formats for presenting risks and risk reductions. Cochrane Database Syst Rev 2011(3):CD006776. [doi: 10.1002/14651858.CD006776.pub2] [Medline: 21412897]
11. Visschers VH, Meertens RM, Passchier WW, de Vries NN. Probability information in risk communication: a review of the research literature. Risk Anal 2009 Feb;29(2):267-287. [doi: 10.1111/j.1539-6924.2008.01137.x] [Medline: 19000070]
12. Bonner C, Jansen J, McKinn S, Irwig L, Doust J, Glasziou P, et al. General practitioners' use of different cardiovascular risk assessment strategies: a qualitative study. Med J Aust 2013 Oct 7;199(7):485-489. [Medline: 24099210]
13. Torley D, Zwar N, Comino EJ, Harris M. GPs' views of absolute cardiovascular risk and its role in primary prevention. Aust Fam Physician 2005 Jun;34(6):503-4, 507 [FREE Full text] [Medline: 15931411]
14. Sposito AC, Ramires JA, Jukema JW, Molina JC, da Silva PM, Ghadanfar MM, et al. Physicians' attitudes and adherence to use of risk scores for primary prevention of cardiovascular disease: cross-sectional survey in three world regions. Curr Med Res Opin 2009 May;25(5):1171-1178. [doi: 10.1185/03007990902846423] [Medline: 19323611]
15. Heeley EL, Peiris DP, Patel AA, Cass A, Weekes A, Morgan C, et al. Cardiovascular risk perception and evidence--practice gaps in Australian general practice (the AusHEART study). Med J Aust 2010 Mar 1;192(5):254-259. [Medline: 20201758]
16. Goldman RE, Parker DR, Eaton CB, Borkan JM, Gramling R, Cover RT, et al. Patients' perceptions of cholesterol, cardiovascular disease risk, and risk communication strategies. Ann Fam Med 2006;4(3):205-212 [FREE Full text] [doi: 10.1370/afm.534] [Medline: 16735521]
17. Grover SA, Lowensteyn I, Joseph L, Kaouache M, Marchand S, Coupal L, Cardiovascular Health Evaluation to Improve Compliance and Knowledge Among Uninformed Patients (CHECK-UP) Study Group. Patient knowledge of coronary risk profile improves the effectiveness of dyslipidemia therapy: the CHECK-UP study: a randomized controlled trial. Arch Intern Med 2007 Nov 26;167(21):2296-2303. [doi: 10.1001/archinte.167.21.2296] [Medline: 18039987]
18. Parkes G, Greenhalgh T, Griffin M, Dent R. Effect on smoking quit rate of telling patients their lung age: the Step2quit randomised controlled trial. BMJ 2008 Mar 15;336(7644):598-600 [FREE Full text] [doi: 10.1136/bmj.39503.582396.25] [Medline: 18326503]
19. Soureti A, Hurling R, Murray P, van Mechelen W, Cobain M. Evaluation of a cardiovascular disease risk assessment tool for the promotion of healthier lifestyles. Eur J Cardiovasc Prev Rehabil 2010 Oct;17(5):519-523. [doi: 10.1097/HJR.0b013e328337ccd3] [Medline: 20195154]
20. National Vascular Disease Prevention Alliance (NVDPA). Guidelines for the management of absolute cardiovascular disease risk. Australia: NVDPA; 2012.
21. National Vascular Disease Prevention Alliance (NVDPA). Guidelines for the assessment of absolute cardiovascular disease risk. Australia: NVDPA; 2009.
22. Bowen GA. Naturalistic inquiry and the saturation concept: a research note. Qualitative Research 2008 Feb 01;8(1):137-152. [doi: 10.1177/1468794107085301]

23. Unilever PLC/Unilever NV. Heart Age Calculator. URL: http://www.heartagecalculator.com [accessed 2012-12-05] [WebCite Cache ID 6PADS5yz3]
24. National Heart Foundation of New Zealand. Your Heart Forecast. URL: http://www.knowyournumbers.co.nz [accessed 2012-12-05] [WebCite Cache ID 6PADOyiss]
25. Kuusela H, Paul P. A comparison of concurrent and retrospective verbal protocol analysis. Am J Psychol 2000;113(3):387-404. [Medline: 10997234]
26. Softonic International SL. SMRecorder 1.2. URL: http://smrecorder.en.softonic.com/ [accessed 2014-04-28] [WebCite Cache ID 6PAEFGX3m]
27. Ritchie J, Spencer L, O'Connor W. Carrying out qualitative analysis. In: Ritchie J, Spencer L, editors. Qualitative Research Practice: A Guide for Social Science Students and Researchers. London: Sage Publications; 2003:219-262.
28. Barbour RS. Checklists for improving rigour in qualitative research: a case of the tail wagging the dog? BMJ 2001 May 5;322(7294):1115-1117 [FREE Full text] [Medline: 11337448]
29. Harle CA, Downs JS, Padman R. Effectiveness of personalized and interactive health risk calculators: a randomized trial. Med Decis Making 2012;32(4):594-605. [doi: 10.1177/0272989X11431736] [Medline: 22247421]
30. Ayers SL, Kronenfeld JJ. Chronic illness and health-seeking information on the Internet. Health (London) 2007 Jul;11(3):327-347. [doi: 10.1177/1363459307077547] [Medline: 17606698]
31. Zikmund-Fisher BJ, Dickson M, Witteman HO. Cool but counterproductive: interactive, Web-based risk communications can backfire. J Med Internet Res 2011;13(3):e60 [FREE Full text] [doi: 10.2196/jmir.1665] [Medline: 21868349]
32. Harle CA, Padman R, Downs JS. Design, implementation, and preliminary evaluation of a web-based health risk calculator. 2009 Presented at: 42nd Hawaii International Conference on System Sciences; 2009; Hawaii. [doi: 10.1109/HICSS.2009.148]
33. Wells S, Kerr A, Eadie S, Wiltshire C, Jackson R. 'Your Heart Forecast': a new approach for describing and communicating cardiovascular risk? Heart 2010 May;96(9):708-713. [doi: 10.1136/hrt.2009.191320] [Medline: 20424153]

Abbreviations

CVD: cardiovascular disease
GP: general practitioner

Edited by G Eysenbach; submitted 16.12.13; peer-reviewed by C Harle, N Mayo; comments to author 10.03.14; revised version received 19.03.14; accepted 14.04.14; published 05.05.14.

Please cite as:
Bonner C, Jansen J, Newell BR, Irwig L, Glasziou P, Doust J, Dhillon H, McCaffery K
I Don't Believe It, But I'd Better Do Something About It: Patient Experiences of Online Heart Age Risk Calculators
J Med Internet Res 2014;16(5):e120
URL: http://www.jmir.org/2014/5/e120/
doi:10.2196/jmir.3190
PMID:24797339

Original Paper

Are People With Chronic Diseases Interested in Using Telehealth? A Cross-Sectional Postal Survey

Louisa Edwards[1], PhD; Clare Thomas[1], PhD; Alison Gregory[1], BSc; Lucy Yardley[2], PhD; Alicia O'Cathain[3], PhD; Alan A Montgomery[4], PhD; Chris Salisbury[1], MD

[1]Centre for Academic Primary Care, School of Social and Community Medicine, University of Bristol, Bristol, United Kingdom

[2]Department of Psychology, University of Southampton, Southampton, United Kingdom

[3]Medical Care Research Unit, School of Health and Related Research (ScHARR), University of Sheffield, Sheffield, United Kingdom

[4]Nottingham Clinical Trials Unit, University of Nottingham, Nottingham, United Kingdom

Corresponding Author:
Louisa Edwards, PhD
Centre for Academic Primary Care
School of Social and Community Medicine
University of Bristol
Canynge Hall
39 Whatley Road
Bristol, BS8 2PS
United Kingdom
Phone: 44 1179287257
Fax: 44 1179287324
Email: Louisa.Edwards@bristol.ac.uk

Abstract

Background: There is growing interest in telehealth—the use of technology to support the remote delivery of health care and promote self-management—as a potential alternative to face-to-face care for patients with chronic diseases. However, little is known about what precipitates interest in the use of telehealth among these patients.

Objective: This survey forms part of a research program to develop and evaluate a telehealth intervention for patients with two exemplar chronic diseases: depression and raised cardiovascular disease (CVD) risk. The survey was designed to explore the key factors that influence interest in using telehealth in these patient groups.

Methods: Thirty-four general practices were recruited from two different regions within England. Practice records were searched for patients with (1) depression (aged 18+ years) or (2) 10-year risk of CVD ≥20% and at least one modifiable risk factor (aged 40-74 years). Within each general practice, 54 patients in each chronic disease group were randomly selected to receive a postal questionnaire. Questions assessed five key constructs: sociodemographics, health needs, difficulties accessing health care, technology-related factors (availability, confidence using technology, perceived benefits and drawbacks of telehealth), and satisfaction with prior use of telehealth. Respondents also rated their interest in using different technologies for telehealth (phone, email and Internet, or social media). Relationships between the key constructs and interest in using the three mediums of telehealth were examined using multivariable regression models.

Results: Of the 3329 patients who were sent a study questionnaire, 44.40% completed it (872/1740, 50.11% CVD risk; 606/1589, 38.14% depression). Overall, there was moderate interest in using phone-based (854/1423, 60.01%) and email/Internet-based (816/1425, 57.26%) telehealth, but very little interest in social media (243/1430, 16.99%). After adjusting for health needs, access difficulties, technology-related factors, and prior use of telehealth, interest in telehealth had largely no association with sociodemographic variables. For both patient groups and for each of the three technology mediums, the most important constructs related to interest in telehealth were having the confidence to use the associated technology, as well as perceiving greater advantages and fewer disadvantages from using telehealth. To illustrate, greater confidence using phone technologies (b=.16, 95% CI 0.002-0.33), while also perceiving more benefits (b=.31, 95% CI 0.21-0.40) and fewer drawbacks (b=-.23, 95% CI -0.28 to -0.17) to using telehealth were associated with more interest in using phone-based telehealth technologies for patients with depression.

Conclusions: There is widespread interest in using phone-based and email/Internet-based telehealth among patients with chronic diseases, regardless of their health status, access difficulties, age, or many other sociodemographic factors. This interest could

be increased by helping patients gain confidence using technologies and through highlighting benefits and addressing concerns about telehealth. While the same pattern exists for social media telehealth, interest in using these technologies is minimal.

(J Med Internet Res 2014;16(5):e123) doi:10.2196/jmir.3257

KEYWORDS

telehealth; Internet; technology; cardiovascular diseases; depression; mental health; chronic disease; survey methodology; patient acceptance of health care

Introduction

Health care systems globally are likely to struggle to cope with the ever-increasing number of people with chronic diseases, and the United Kingdom is no exception [1,2]. For example, nearly a third of the population has a chronic disease, and this figure is projected to rise by 23% within 25 years [3]. Furthermore, health service use is high among this group of patients. Patients with chronic diseases use a large portion of general practitioner (52%) and outpatient (65%) appointments, and an estimated 69% of the primary and acute care budget is spent supporting patients with chronic diseases [3]. Therefore, exploring alternative ways of delivering care, supporting patients, and managing chronic diseases is needed, particularly in light of the financial pressures currently facing health care systems.

There is considerable international interest in telehealth as a possible alternative to face-to-face care for people with chronic diseases [2,4]. Similar to the World Health Organization [5], we define telehealth as the use of technology to support the remote delivery of health care and promote self-management. Both cost-effectiveness and clinical effectiveness should be demonstrated before telehealth becomes a mainstay in the health care system. Although some pilot studies have suggested large potential cost savings [6], a recent large randomized controlled trial of telehealth and telecare suggested that it was unlikely to be cost-effective [7]. The results from this Whole System Demonstrator trial did, however, suggest reduced mortality and emergency admission rates [8]. While this latter result is promising, some reviews conclude that the evidence for the clinical effectiveness of telehealth is, in fact, mixed [9]. It may be that some forms of telehealth can lead to improvements in patients with certain chronic diseases, such as cardiovascular disease (CVD) [10] and depression [11], since there is evidence for the effectiveness of some specific telehealth interventions in these conditions. For example, blood pressure self-monitoring, combined with self-titration of antihypertensive medication resulted in significant reductions in systolic blood pressure compared to usual care [12], while a therapist-delivered online cognitive behavioral therapy program led to greater recovery from depression than usual care [13].

Equity, as well as cost-effectiveness, is an important consideration for health care systems. Telehealth has the potential to improve care for patients with difficulty accessing traditional services, such as those who are housebound or live in rural areas [14]. These patients are also likely to be those who have the greatest health needs [15]. Additionally, since telehealth can enable a patient to monitor their own vital signs at home (eg, blood pressure), it may be more convenient and comfortable, enhance independence, and empower patients [16].

To realize the benefits of telehealth, patients must engage with and make use of it [17]. Some previous studies have suggested limited engagement with telehealth interventions in patients with chronic diseases [18], and a refusal rate of up to 75% from those invited to join telehealth trials [19]. If telehealth is to make an important contribution to the health care system for managing chronic diseases, it is imperative to identify, and then appropriately target, the factors that influence interest in telehealth, because people must be interested if they are going to make use of it. A systematic review of 52 studies on patient acceptance of computer-based health information technologies concluded that the majority of literature to date has focused on patient factors, such as sociodemographic variables [20]. For example, some previous research has suggested that interest in telehealth is highest in younger, educated, and affluent patients [21,22], but these characteristics are inversely associated with the prevalence of chronic diseases [3]. A recent review commissioned by the National Health Service (NHS) in England [22] identified five categories of barriers and facilitators to telehealth services: user characteristics, technological aspects, characteristics of services, social aspects of use, and telehealth services in use. However, both this and the aforementioned review [20] were not limited to patients with chronic diseases, nor did they aim to quantitatively assess the relative importance of factors influencing interest in telehealth. Nonetheless, in line with some of the findings from these reviews, we reasoned that both structural and evaluative factors would be key influences of interest in telehealth; namely, whether these patients have the technology readily available to use, their confidence in using technology, and their attitude towards telehealth. Moreover, if those with the greatest health needs and greatest difficulties in accessing health care are indeed interested in using telehealth, a large gap in unmet need could be filled.

We carried out a study to investigate the factors that influence interest in telehealth among patients with chronic diseases. This work was conducted as part of a larger research program exploring the potential role of an existing health service in England, NHS Direct, in providing support for chronic diseases via the telephone and Internet. For this reason, we did not name specific or existing telehealth services but asked a large number of respondents about their interest in using several types of technology that *could* be used for telehealth. The aim of the current study was to determine whether interest in telehealth among patients with chronic diseases is related to health needs, difficulties in accessing health care, or technology-related factors, including availability and attitudes to technology, while

also considering the role of sociodemographic factors and taking into account prior experience using NHS Direct.

Methods

Design

We used a cross-sectional postal survey for this study.

Choice of Chronic Diseases

This study focused on two exemplar chronic diseases. The first was risk of a cardiovascular event (heart attack or stroke) ≥20% over the next 10 years. This approach recognizes that hypertension, obesity, and hyperlipidemia are risk factors for CVD, rather than conditions, and it is more appropriate to consider raised CVD risk as a chronic disease [23]. The second exemplar chronic disease was depression. These two conditions were chosen to represent different types of chronic diseases, both of which are common, in which there is considerable unmet need, and where there is some evidence that particular forms of telehealth may be of clinical benefit [10-13].

Sampling and Recruitment

General practices in two geographical areas of the United Kingdom, the south west and the north east, were invited to take part in the study. General practices were intentionally selected to represent a wide mix of socioeconomic characteristics of patients. Between August 2010 and May 2011, a query was run on practice records to identify patients with either depression (aged ≥18 years, had consulted their doctor about a mental health issue, and were prescribed an antidepressant medication within the last year) or raised risk of CVD (aged 40-74 years, QRISK2 [24] or Framingham [25] 10-year risk ≥20%, and at least one modifiable risk factor, including hypertension, obesity, or smoking). We calculated QRISK2 to assess CVD risk where possible, but since this score was not available through all general practice computer systems, we used Framingham risk scores in some practices. Patients were excluded if they were terminally ill, had cognitive impairment, or had a severe mental health condition, such as psychosis.

Fifty-four patients per practice from each of the two groups of eligible patients were selected using stratified random sampling. We sampled females and males in proportion to the number of eligible patients in each general practice. The CVD risk group was further stratified by age, such that equal proportions of young (aged 40-59 years) and older (60-74 years) participants were selected. This was because CVD risk ≥20% is more prevalent among older individuals, while access to technology is inversely associated with age [21]. Prior to invitation, general practitioners (GPs) reviewed the patient lists and excluded any patients for whom it would be inappropriate to send a questionnaire (eg, due to recent bereavement). The remaining patients were then mailed a letter by their general practice inviting them to take part in a study looking at new ways the NHS could help people to improve their health, as well as a participant information pamphlet and a questionnaire. Patients were asked to return a blank questionnaire if they did not want to take part. Those who did not respond were sent up to two postal reminders at approximately 2-week intervals. All correspondence was sent by staff at the patient's general practice, and the researchers did not have access to patient identifiable data at any point. Ethical approval was granted by the Southmead Research Ethics Committee.

Sample Size

Assuming an approximate 60% response rate, inviting 54 patients from each of 32 practices would provide around 960 respondents for each chronic disease group. This would provide 80% power to detect an absolute difference of ≤9.2% points in interest in using telehealth (binary outcome, equivalent odds ratio ≤1.45), with two-sided 5% alpha.

Measures

The questionnaire included questions about the key constructs that we hypothesized would predict interest in telehealth; namely, sociodemographics, health needs, difficulties accessing health care, availability and attitudes to technology, and prior use of telehealth. In order to ensure the coherence of the questions included to assess these constructs and to reduce the questionnaire items to a smaller number of factors for data analysis, principal components analyses (PCA) with orthogonal (varimax) rotation were carried out using STATA on constructed items. Decisions regarding the number of factors to extract were based on Kaiser's criterion (eigenvalues >1.0), by examining the scree plot and the subjective coherence of the factors. For each factor, items with an association ≥.3 were retained [26]. Next, the reliability of each factor was examined with Cronbach alpha, whereby coefficients above .70 indicate adequate reliability. Finally, mean summary scores for each reliable factor were calculated for individuals providing ratings for ≥50% of the relevant items. We treated each factor as a scale and labeled it according to the questions it comprised.

Outcome Variable

Interest in telehealth was assessed using questions about the participant's interest in using a range of technologies. The item reduction techniques described above resulted in three summary scores for interest in telehealth, which relate to interest in three types of technology: Phone (alpha=.82; landline or mobile phone), Email/Internet (alpha=.94; using email or doing searches on the Internet), and Social Media (alpha=.85; using chat rooms and social networking sites). These "interest" summary scores were equal to the averaged sum of responses to three question items each (range: 1-3), such that each corresponding summary score ranged from 1.0 ("Not at all interested"/"I don't know what this is") to 3.0 ("Very interested"), with scores of 2.0 equal to "Fairly interested".

Explanatory Variables

Questions about sociodemographic characteristics of the respondents included sex, ethnicity [27], age group, employment status [28], educational qualifications [29], and home ownership [30]. These questions were based on those used in previous validated surveys where possible.

Health needs were assessed using the Short-Form (SF-12v2) Health Survey, version 2 [31]. Physical Component Summary (PCS) and Mental Component Summary (MCS) scores were derived from the 12 items using proprietary scoring software

(QualityMetric, Incorporated). These indexes of physical and mental health functioning are standardized with a mean of 50 and standard deviation of 10, such that lower scores indicate poorer health or greater needs.

The remainder of the questionnaire contained items constructed for the purposes of this research, although guided and informed by relevant literature, and piloted with service users in advance. These are described below, while the specific questions that comprise each scale are located in Multimedia Appendix 1.

Difficulty accessing health care was assessed using a series of questions that were based on themes identified through previous research [32,33]. Two "access difficulty" summary scores resulted: Service Delivery difficulties (7 items, alpha=.87) included questions about the convenience of accessing health care, as well as the nature or quality of the care itself (eg, getting the right amount of care), and Physical Access difficulties (4 items, alpha=.78) included questions about trouble getting to appointments due to physical, psychological, and transport problems, including cost. These summary scores ranged between 1.0 ("No difficulty") and 3.0 ("Lots of difficulty").

Technology-related factors included questions about availability of technologies and attitudes towards telehealth. Technology availability was assessed by asking respondents which of a range of technologies were easily available for them to use. Phone Availability (2 items: landline, mobile) and Email/Internet Availability (2 items: have email address, Internet access) scores were formed by summing tallies (0=Absent, 1=Present) for these technologies.

Questions about attitudes towards telehealth were based on the theory of planned behavior [34]. This theory suggests that perceived behavioral control—a concept capturing the extent to which one believes one is able to perform a behavior—directly influences one's intention to carry out a behavior and may predict behavior itself. Beliefs about one's capability, which should be reflected in confidence levels, affect perceived behavioral control. Therefore, questions about confidence using different types of technology were devised. After item reduction, there were three clusters representing confidence in using Phone-based technologies (3 items, alpha=.74), Email/Internet-based technologies (3 items, alpha=.96), and Social Media-based technologies (3 items, alpha=.88). Again, larger scores indicate greater technology confidence [range: 1.0 ("Not at all confident"/"I have never tried this"/"I don't know what this is") to 3.0 ("Extremely confident")].

The theory of planned behavior also states that positive or negative attitudes towards a behavior predict one's intention to perform that behavior and are influenced by beliefs about the advantages and disadvantages of that behavior [34]. Hence, items about the potential advantages and disadvantages of telehealth were generated based on previous qualitative research [35]. Summary scores for telehealth Advantages (7 items, alpha=.87) and Disadvantages (7 items, alpha=.90) were similarly formed (range: 1.0 ("Strongly disagree") to 5.0 ("Strongly agree"); higher scores reflect greater perceived advantages and disadvantages.

Finally, satisfaction with prior use of telehealth that was delivered by NHS Direct was evaluated by a single item. Respondents rated how satisfied they were with previous use of NHS Direct services on a scale from 1 ("Not at all") to 5 ("Extremely"). NHS Direct is a service freely available throughout England that provides health assessments, information, and advice. It currently provides telehealth via telephone and its interactive website, but NHS Direct and other similar services could act as a provider of a wider range of telehealth services.

Statistical Analysis

The primary analysis investigated the extent to which interest in the use of telehealth was related to five key constructs: sociodemographic factors, health needs (including physical and mental health), access difficulties (including service delivery and physical access), technology-related factors (availability of technology and attitudes towards telehealth), and satisfaction with prior use of telehealth. We first used appropriate descriptive statistics (mean and SD, or n and %) to summarize the sociodemographic characteristics of respondents, and their needs, access difficulties, technology factors, and interest in using telehealth. This included an exploration of how needs, access, and technology factors varied by age and chronic disease group. We then used multivariable regression models to examine associations between these variables and interest in telehealth, adjusting for the other variables in the model, and taking into account the stratified survey design.

Results

Response Rate

Thirty-four general practices took part in the survey. GPs excluded 5.23% of the CVD risk group (96/1836) and 11.23% of patients with depression (201/1790) prior to mailing questionnaires. Of the 3329 patients who were sent a study questionnaire, 1478 (44.40%) returned it. The response rate was higher for patients with CVD risk (872/1740, 50.11%) than for depression (606/1589, 38.14%). Separate logistic regression analyses for the two patient groups revealed that response rates for both CVD risk and depression were higher in older people, whereas the likelihood of responding did not differ by respondent sex or location (Table 1).

Table 1. Demographic differences between responders and non-responders by patient group.

Characteristics	Patient group: CVD risk (n=1635)			Patient group: depression (n=1497)		
		Responded (n=828)	No response (n=807)		Responded (n=583)	No response (n=914)
	OR (95% CI)	n (%)	n (%)	OR (95% CI)	n (%)	n (%)
Age (years)						
18-29	–	–	–	Referent	64 (11.0)	218 (23.9)
30-44	Referent[a]	18 (2.2)	39 (4.8)	2.0 (1.4-2.8)	166 (28.5)	311 (34.0)
45-59	1.6 (0.6-3.9)	290 (35.0)	391 (48.5)	3.5 (2.5-5.0)	197 (33.8)	232 (25.4)
60-74	3.0 (1.2-7.2)	514 (62.1)	377 (46.7)	4.4 (3.0-6.4)	112 (19.2)	90 (9.8)
75+	–	6 (0.7)	0 (0.0)	2.8 (1.6-5.0)	44 (7.5)	63 (6.9)
Sex [b]	1.3 (0.9-1.7)			1.4 (0.9-1.9)		
Male		620 (74.9)	621 (77.0)		148 (25.4)	295 (32.3)
Female		208 (25.1)	186 (23.0)		435 (74.6)	619 (67.7)
Location [c]	0.8 (0.6-1.0)			1.0 (0.8-1.4)		
Bristol		438 (52.9)	386 (47.8)		282 (48.4)	472 (51.6)
Sheffield		390 (47.1)	421 (52.2)		301 (51.6)	442 (48.4)

[a]Referent age group is 40-44 years.

[b]Sex (0=Male, 1=Female).

[c]Location (0=Bristol, 1=Sheffield).

Sample Characteristics

Patients with CVD risk (mean 61.9 years, SD 7.8) were older than those with depression (mean 49.1 years, SD 15.9), reflecting the inclusion criteria. Three-quarters of the CVD risk group were male (654/872, 75.0%), while three-quarters of the depression group were female (452/606, 74.6%). Both patient groups were predominantly Caucasian (CVD 825/851, 96.9%; depression 575/594, 96.8%), most were not currently employed (unemployed, studying, retired, etc) (CVD 498/861, 57.8%; depression 317/597, 53.1%), only a minority had higher education (CVD 212/872, 24.3%; depression 222/606, 36.6%), while the majority were home owners (CVD 647/859, 75.3%; depression 410/595, 68.9%).

Overview of Health Needs, Access Difficulties, and Technology-Related Factors

As expected, patients with CVD risk reported poorer physical than mental health, whereas the reverse was true of patients with depression (Table 2). While the reported physical health of patients with CVD risk was half a standard deviation below that of the national average (UK mean 50.9, SD 9.4), the reported mental health of patients with depression was more than 1.5 standard deviations below this average (UK mean 50.9, SD 9.4) [36].

Few patients reported access difficulties, with all summary scores approximating the “No difficulty” response category (Table 2). Despite these low mean summary scores, an important minority of participants indicated some difficulty in accessing health care, and both patient groups were more likely to report having service delivery than physical access difficulties. For example, 27.86% (399/1432) of patients reported difficulties getting care when they need it most (service delivery), while 14.23% (206/1448) reported difficulties traveling to appointments due to physical health (physical access).

Technology availability was high across both patient groups (Table 2). Phone technologies were more prevalent than computer-based technologies and markedly so for the patients with CVD risk. In fact, nearly all patients had access to phone technologies. Across patient groups, age was associated only with availability of the computer-based technologies: 90.0% (115/128) of the two youngest age groups (18-44 years), 78.1% (400/512) of those aged 45-59 years, and 60.5% (393/650) of those aged 60-74 years reported that they have these technologies readily available to use. It is among only the oldest, and proportionally smallest, age group (n=49) that less than half the respondents (13, 26.5%) report easy access to computer technologies.

Table 2. Health needs, access difficulties, technology-related factors, and satisfaction with prior telehealth use by patient group.

Characteristics	Patient group: CVD risk	Patient group: depression
Health needs, mean (SD), n		
PCS	45.3 (11.8), 777	47.3 (13.8), 547
MCS	49.8 (10.5), 777	37.7 (12.9), 547
Access difficulties, mean (SD), n [a]		
Service delivery difficulties	1.3 (0.4), 848	1.5 (0.5), 595
Physical access difficulties	1.1 (0.3), 854	1.2 (0.4), 594
Technology-related factors		
Phone availability, % (n)[b]	98.4 (855/869)	99.3 (595/599)
Email/Internet availability, % (n)[b]	67.2 (584/869)	80.3 (481/599)
Phone confidence, mean (SD), n[c]	2.5 (0.6), 861	2.5 (0.6), 596
Email/Internet confidence, mean (SD), n[c]	2.0 (0.9), 851	2.3 (0.8), 595
Social media confidence, mean (SD), n[c]	1.3 (0.6), 847	1.6 (0.8), 594
Telehealth advantages, mean (SD), n[d]	3.6 (0.8), 853	3.7 (0.7), 588
Telehealth disadvantages, mean (SD), n[d]	3.5 (0.9), 860	3.3 (0.9), 593
Satisfaction with prior telehealth use, mean (SD), n [e]		
NHS Direct satisfaction	3.4 (1.2), 247	3.4 (1.2), 336

[a]Range: 1.0-3.0, where higher scores indicate greater access difficulties. Service delivery difficulties included questions about the convenience of accessing health care, as well as the nature or quality of the care itself, eg, getting the right amount of care. Physical access difficulties included questions about trouble getting to appointments due to physical, psychological, and transport problems, including cost, as detailed in the Methods section.

[b]Technology availability includes having one or more forms of relevant technology.

[c]Range: 1.0-3.0, where higher scores indicate greater technology confidence.

[d]Range: 1.0-5.0, where higher scores indicate greater perceived advantages and disadvantages of telehealth.

[e]Range: 1.0-5.0, where higher scores indicate greater satisfaction with past NHS Direct use.

Technology confidence ratings were similar between patient groups, but they varied somewhat across the technology types (Table 2) and age groups (Figure 1). In general, patients reported greatest confidence using phone technologies, with mean summary scores approaching the "Extremely confident" response category, and least confidence using the social media technologies, with mean summary scores close to "Not at all confident". Respondents were "Quite confident" with the email-based and Internet-based technology category.

Figure 1 shows the proportion of depression respondents reporting they were "Extremely confident" or "Quite confident" using the various technologies across the different age groups. The pattern of findings was similar among the CVD risk group (data not shown). Least associated with age were the phone technologies, which received high confidence ratings by all age groups. The one exception was low confidence in text messaging by the oldest age group. Although confidence using email/Internet and social media technologies consistently decreased with age, more than half of the respondents in all age groups (except the over-75s) reported confidence using email/Internet technologies. Conversely, confidence using social media technologies was strongly related to age, with only the younger age groups expressing confidence.

Summary scores indicate similar levels of perceived advantages and disadvantages of using telehealth across patient groups (Table 2). The most highly endorsed advantages were convenience and ability of telehealth to be delivered when and where one desires (Table 3). Dislike of non–face-to-face care and concerns over security issues emerged as the top disadvantages of telehealth (Table 3).

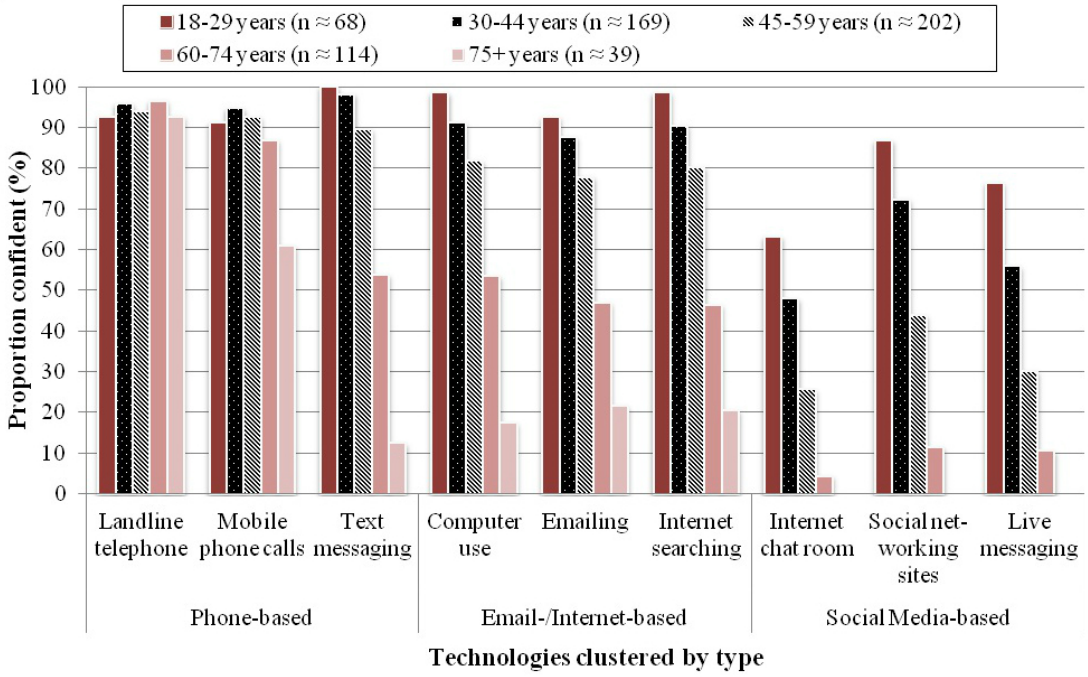
Of those respondents that had ever used NHS Direct (Table 2), the majority were satisfied with that experience: 26.9% (157/583) were "Moderately", 33.1% (193/583) were "Quite a bit", and 18.9% (110/583) were "Extremely" satisfied.

Table 3. Proportion of respondents agreeing[a] with each of the potential advantages and disadvantages of using telehealth by patient group.

	Patient group: CVD risk, % (n)	Patient group: depression, % (n)
Advantages		
I would like being able to choose to get support at times that are best for me.	81.7 (696/852)	87.4 (514/588)
I would find it reassuring to be able to get support when I feel that I need it most.	81.0 (689/851)	85.7 (504/588)
I would like being able to get support in my own home.	67.7 (573/846)	71.2 (418/587)
Getting support with my health by phone or computer would be valuable to me.	54.9 (466/849)	60.8 (360/592)
I could save money by not having to travel to appointments.	50.7 (426/840)	51.4 (299/582)
Getting support in this way would help me to feel more independent.	48.8 (413/847)	54.2 (318/587)
It would make me feel special to be getting 'extra' support when I feel that I need it most.	41.9 (354/844)	42.5 (247/581)
Disadvantages		
I would dislike being unable to see the person face-to-face.	66.6 (571/858)	60.2 (357/593)
I would be concerned about the security of the information that I give.	63.3 (544/860)	60.3 (357/592)
I would not want to discuss sensitive issues over the phone or using a computer	61.9 (532/859)	54.9 (325/592)
I would dislike speaking to someone other than a doctor about my health.	53.7 (462/860)	45.8 (271/592)
I would worry about relying too much on the technology.	52.3 (447/854)	42.2 (247/586)
I would worry about the possibility of the equipment not working.	45.4 (387/852)	37.8 (222/588)
Getting support in this way would make me feel anxious about my health.	33.2 (284/855)	26.0 (153/588)

[a]A response of either "Strongly agree" or "Agree" on a 5-point scale was considered agreement.

Figure 1. Proportion of depression respondents within each age group reporting confidence using individual technologies.



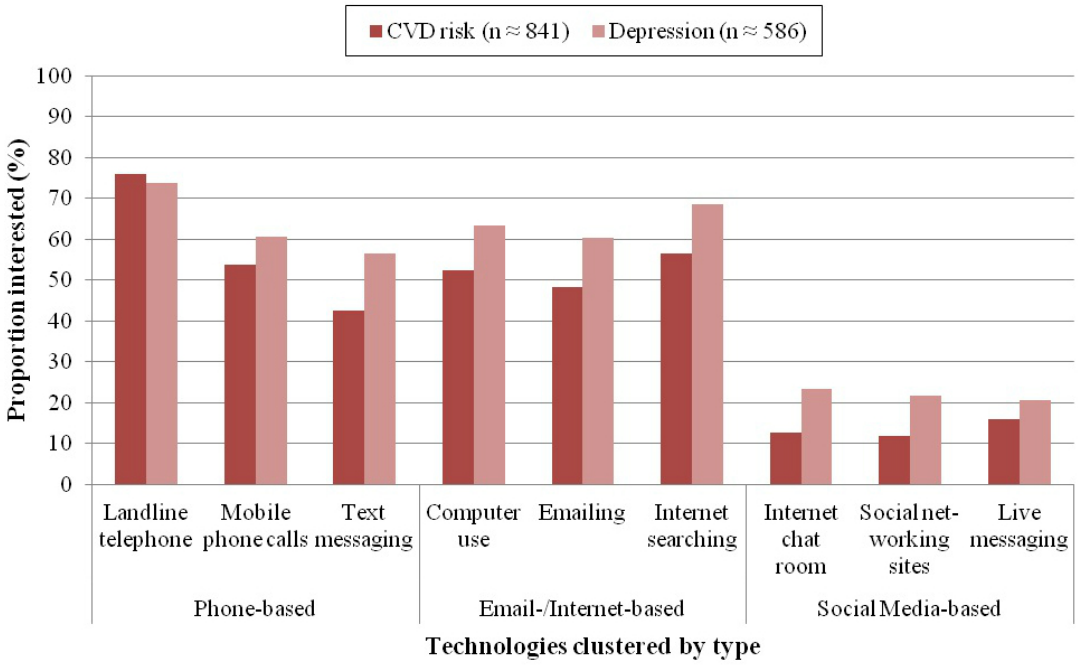
Overview of Interest in Using Telehealth

Regardless of patient group, there was moderate interest in phone (CVD risk mean 1.7, SD 0.6; depression mean 1.9, SD 0.7) and email/Internet technologies (CVD risk mean 1.7, SD 0.7; depression mean 1.9, SD 0.7). These mean summary scores approximate the "Fairly interested" response category. In contrast, there was very little interest in the social media technologies (CVD risk mean 1.2, SD 0.4; depression mean 1.3, SD 0.5).

Figure 2 shows which individual technologies respondents were more or less interested in using, with ratings of interest dichotomized into some versus no interest for ease of interpretation. This shows that patients with depression were more interested than those with CVD risk in nearly every form of technology for telehealth. There was a clear preference for

the landline telephone (1072/1428, 75.07% overall), followed by finding information on the Internet (876/1427, 61.39% overall). Again, there was hardly any interest in the social media technologies. Averaging across the technology types and both patient groups (CVD risk and depression), there was moderate interest in using phone-based (854/1423, 60.01%) and email/Internet-based (816/1425, 57.26%) telehealth, but very little interest in social media (243/1430, 16.99%).

Figure 2. Proportion of respondents interested in individual telehealth technologies by patient group.



What Factors are Associated With Interest in Telehealth?

To address the main research question, sociodemographics, health needs, access difficulties, technology-related factors, and satisfaction with prior telehealth use were simultaneously regressed on interest for each of the three telehealth mediums—phone-based, email/Internet-based, and social media–based telehealth—separately for each patient group (CVD risk and depression). From these multivariable linear regression analyses, three variables were reliably related to interest in telehealth: greater technology confidence, and perceiving both greater advantages and fewer disadvantages of telehealth were associated with more interest in using telehealth. Moreover, these factors were consistently related to interest in each of the three telehealth mediums for both patient groups (Tables 4-6). Importantly, however, the technology confidence finding is modality-specific. This means that greater phone confidence is associated with greater phone-based telehealth interest, greater email/Internet confidence is associated with greater email/Internet-based telehealth interest, while greater confidence using social media technologies is associated with greater interest in social media–based telehealth. Therefore, while confidence using a particular type of technology discriminated between interest in different modes of telehealth, perceiving greater benefits of and fewer drawbacks to using telehealth was uniformly related to greater interest.

Three other consistent findings emerged. First, for patients with depression but not those with CVD risk, greater difficulties with getting convenient, high-quality care (service delivery aspects of access) were related to more interest in phone-based and email/Internet-based telehealth technologies. Second, as anticipated, greater satisfaction with previous use of NHS Direct was associated with heightened interest in future use of phone-based telehealth, and this was consistent across both patient groups. Third, there was more interest in email/Internet-based and social media–based telehealth among those with CVD risk who were not home owners. Apart from these findings, the remaining variables in the model were unimportant to telehealth interest. After adjusting for the other variables in the model, health needs, access difficulties, technology availability, and even sociodemographic factors did not reliably and consistently have an independent effect on interest in telehealth, either across patient groups or across more than one telehealth type within a patient group.

Table 4. Multivariable analysis of factors associated with interest in phone-based telehealth[a].

Characteristics		Patient group: CVD risk (n=676)		Patient group: depression (n=489)	
		b (95% CI)	*P*	b (95% CI)	*P*
Sociodemographic factors					
Age group (years)					
	18-29	–	–	Referent	
	30-44	Referent (40-44 years)		-.054 (-0.187 to 0.079)	
	45-59	-.101 (-0.379 to 0.178)		-.016 (-0.201 to 0.168)	
	60-74	-.122 (-0.421 to 0.176)		-.028 (-0.178 to 0.122)	
	75+	.301 (-0.484 to 1.086)	.35[b]	.087 (-0.303 to 0.476)	.79[b]
Sex[c]		-.018 (-0.122 to 0.086)	.72	-.104 (-0.226 to 0.018)	.09
Ethnicity[d]		-.061 (-0.312 to 0.191)	.63	.161 (-0.076 to 0.398)	.18
Employed[e]		.108 (0.015 to 0.201)	.02	.049 (-0.056 to 0.155)	.35
Higher education[f]		-.009 (-0.092 to 0.074)	.83	-.157 (-0.272 to -0.042)	.01
Home owner[g]		-.088 (-0.206 to 0.030)	.14	-.173 (-0.281 to -0.065)	.003
Location[h]		-.017 (-0.123 to 0.090)	.75	-.072 (-0.170 to 0.026)	.14
Health needs					
PCS		.001 (-0.003 to 0.005)	.67	.005 (0.001 to 0.009)	.02
MCS		-.002 (-0.006 to 0.001)	.21	.004 (0.001 to 0.008)	.02
Access difficulties [i]					
Service delivery		-.093 (-0.232 to 0.046)	.19	.205 (0.069 to 0.340)	.004
Physical access		.185 (-0.010 to 0.379)	.06	-.066 (-0.280 to 0.148)	.53
Technology-related factors					
Phone availability[j]		.107 (-0.039 to 0.254)	.15	.203 (0.025 to 0.382)	.03
Email/Internet availability[j]		-.012 (-0.118 to 0.095)	.82	-.089 (-0.178 to 0.0003)	.05
Phone confidence[i]		.254 (0.151 to 0.358)	<.001	.164 (0.002 to 0.326)	.048
Email/Internet confidence[i]		-.075 (-0.197 to 0.046)	.22	-.011 (-0.111 to 0.088)	.82
Social media confidence[i]		.065 (-0.018 to 0.147)	.12	-.025 (-0.110 to 0.060)	.56
Telehealth advantages[i]		.296 (0.240 to 0.352)	<.001	.308 (0.213 to 0.404)	<.001
Telehealth disadvantages[i]		-.201 (-0.261 to -0.140)	<.001	-.226 (-0.282 to -0.170)	<.001
Past telehealth satisfaction [i]					
NHS Direct		.088 (0.025 to 0.151)	.01	.046 (0.001 to 0.090)	.045

[a]The associations have been adjusted by all other variables in this fully adjusted model, and the stratified survey design has been taken into account in the analysis. Interest in phone-based telehealth scores range from 1.0-3.0, with higher scores indicating more interest.

[b]Indicates *P* value from Wald test.

[c]Sex (0=Male, 1=Female).

[d]Ethnicity (0=Non-Caucasian, 1=Caucasian).

[e]Employed (0=Not employed, 1=Employed).

[f]Higher Education (0=No higher education, 1=Some higher education).

[g]Home Owner (0=Non-home owner, 1=Home owner).

[h]Location (0=Bristol, 1=Sheffield).

[i]Higher scores indicate greater access difficulties, technology confidence, advantages, disadvantages, and satisfaction.

[j]Technology Availability (0=Not available, 1=One form available, 2=Both available).

Table 5. Multivariable analysis of factors associated with interest in email/Internet-based telehealth[a].

Characteristics		Patient group: CVD risk (n=681)		Patient group: depression (n=488)	
		b (95% CI)	*P*	b (95% CI)	*P*
Sociodemographic factors					
Age group (years)					
	18-29	–	–	Referent	
	30-44	Referent (40-44 years)		-.041 (-0.278 to 0.197)	
	45-59	-.184 (-0.373 to 0.005)		.006 (-0.198 to 0.209)	
	60-74	-.132 (-0.346 to 0.082)		.045 (-0.160 to 0.250)	
	75+	.371 (-0.120 to 0.863)	.01[b]	.032 (-0.299 to 0.364)	.89[b]
Sex[c]		-.035 (-0.122 to 0.052)	.42	.008 (-0.114 to 0.130)	.90
Ethnicity[d]		.071 (-0.133 to 0.276)	.48	-.042 (-0.221 to 0.138)	.64
Employed[e]		.086 (-0.010 to 0.182)	.08	-.048 (-0.146 to 0.049)	.32
Higher education[f]		.070 (-0.068 to 0.207)	.31	-.037 (-0.155 to 0.080)	.52
Home owner[g]		-.148 (-0.266 to -0.031)	.02	-.059 (-0.158 to 0.040)	.24
Location[h]		-.002 (-0.096 to 0.092)	.97	-.090 (-0.164 to -0.016)	.02
Health needs					
PCS		-.002 (-0.007 to 0.003)	.44	.003 (-0.001 to 0.008)	.16
MCS		-.006 (-0.013 to 0.001)	.08	-.005 (-0.010 to 0.0004)	.07
Access difficulties [i]					
Service delivery		-.013 (-0.124 to 0.098)	.81	.087 (0.005 to 0.168)	.04
Physical access		.017 (-0.249 to 0.284)	.90	-.127 (-0.240 to -0.013)	.03
Technology-related factors					
Phone availability[j]		-.099 (-0.211 to 0.012)	.08	.093 (-0.095 to 0.282)	.32
Email/Internet availability[j]		.158 (0.091 to 0.225)	<.001	.101 (-0.013 to 0.215)	.08
Phone confidence[i]		-.021 (-0.121 to 0.079)	.68	-.199 (-0.367 to -0.032)	.02
Email/Internet confidence[i]		.304 (0.219 to 0.389)	<.001	.403 (0.295 to 0.512)	<.001
Social media confidence[i]		-.002 (-0.157 to 0.154)	.98	.003 (-0.075 to 0.080)	.95
Telehealth advantages[i]		.226 (0.165 to 0.286)	<.001	.237 (0.150 to 0.324)	<.001
Telehealth disadvantages[i]		-.244 (-0.310 to -0.179)	<.001	-.211 (-0.287 to -0.134)	<.001
Past telehealth satisfaction [i]					
NHS Direct		.029 (-0.036 to 0.094)	.38	.040 (-0.034 to 0.115)	.28

[a]The associations have been adjusted by all other variables in this fully adjusted model, and the stratified survey design has been taken into account in the analysis. Interest in email/Internet-based telehealth scores range from 1.0-3.0, with higher scores indicating more interest.

[b]Indicates *P* value from Wald test.

[c]Sex (0=Male, 1=Female).

[d]Ethnicity (0=Non-Caucasian, 1=Caucasian).

[e]Employed (0=Not employed, 1=Employed).

[f]Higher Education (0=No higher education, 1=Some higher education).

[g]Home Owner (0=Non-home owner, 1=Home owner).

[h]Location (0=Bristol, 1=Sheffield).

[i]Higher scores indicate greater access difficulties, technology confidence, advantages, disadvantages, and satisfaction.

[j]Technology Availability (0=Not available, 1=One form available, 2=Both available).

Table 6. Multivariable analysis of factors associated with interest in social media–based telehealth[a].

Characteristics		Patient group: CVD risk (n=680)		Patient group: depression (n=489)	
		b (95% CI)	*P*	b (95% CI)	*P*
Sociodemographic factors					
Age group (years)					
	18-29	–	–	Referent	
	30-44	Referent (40-44 years)		.146 (-0.087 to 0.379)	
	45-59	-.092 (-0.301 to 0.117)		.187 (-0.036 to 0.411)	
	60-74	-.030 (-0.246 to 0.186)		.174 (-0.033 to 0.382)	
	75+	.047 (-0.286 to 0.380)	.25[b]	.061 (-0.163 to 0.285)	.31[b]
Sex[c]		-.025 (-0.080 to 0.030)	.36	.029 (-0.061 to 0.118)	.52
Ethnicity[d]		-.046 (-0.255 to 0.164)	.66	.036 (-0.208 to 0.279)	.77
Employed[e]		.040 (-0.031 to 0.110)	.26	-.007 (-0.114 to 0.100)	.90
Higher eEducation[f]		.018 (-0.057 to 0.093)	.63	-.033 (-0.147 to 0.080)	.55
Home owner[g]		-.118 (-0.198 to -0.038)	.01	-.097 (-0.232 to 0.038)	.15
Location[h]		-.015 (-0.061 to 0.032)	.53	-.005 (-0.082 to 0.073)	.91
Health needs					
PCS		-.001 (-0.004 to 0.001)	.28	-.001 (-0.006 to 0.004)	.67
MCS		-.002 (-0.007 to 0.003)	.44	-.001 (-0.007 to 0.004)	.63
Access difficulties [i]					
Service delivery		-.079 (-0.161 to 0.002)	.06	.029 (-0.098 to 0.156)	.64
Physical access		.029 (-0.118 to 0.176)	.70	-.070 (-0.200 to 0.060)	.28
Technology-related factors					
Phone availability[j]		-.109 (-0.256 to 0.039)	.14	.077 (-0.096 to 0.250)	.37
Email/Internet availability[j]		.016 (-0.029 to 0.062)	.48	.013 (-0.073 to 0.099)	.76
Phone confidence[i]		.012 (-0.071 to 0.095)	.78	-.068 (-0.218 to 0.082)	.36
Email/Internet confidence[i]		.001 (-0.056 to 0.057)	.98	-.038 (-0.118 to 0.041)	.33
Social media confidence[i]		.243 (0.132 to 0.355)	<.001	.361 (0.282 to 0.441)	<.001
Telehealth advantages[i]		.096 (0.045 to 0.146)	.001	.176 (0.106 to 0.245)	<.001
Telehealth disadvantages[i]		-.072 (-0.128 to -0.016)	.01	-.123 (-0.191 to -0.054)	.001
Past telehealth satisfaction [i]					
NHS Direct		.033 (-0.006 to 0.072)	.10	-.006 (-0.051 to 0.040)	.80

[a]The associations have been adjusted by all other variables in this fully adjusted model, and the stratified survey design has been taken into account in the analysis. Interest in social media–based telehealth scores range from 1.0-3.0, with higher scores indicating more interest.

[b]Indicates *P* value from Wald test.

[c]Sex (0=Male, 1=Female).

[d]Ethnicity (0=Non-Caucasian, 1=Caucasian).

[e]Employed (0=Not employed, 1=Employed).

[f]Higher Education (0=No higher education, 1=Some higher education).

[g]Home Owner (0=Non-home owner, 1=Home owner).

[h]Location (0=Bristol, 1=Sheffield).

[i]Higher scores indicate greater access difficulties, technology confidence, advantages, disadvantages, and satisfaction.

[j]Technology Availability (0=Not available, 1=One form available, 2=Both available).

Discussion

Principal Findings

Patients with two very different chronic diseases are interested in using phone-based and email or Internet-based telehealth, but not in telehealth via social media websites. Interest in all three forms of telehealth appears to stem from the perceived advantages and disadvantages of telehealth, as well as confidence using the relevant technology. This is significant because beliefs and levels of confidence are far more malleable than most of the other constructs included in this study, such as technology availability or socioeconomic status. It is also noteworthy that these other constructs were not consistently and independently associated with interest in telehealth. First, interest in telehealth was not reliably related to health needs. This suggests that willingness to use telehealth spans across those with good and poor health. Furthermore, it is not only those who have difficulty accessing traditional health care who are motivated to use telehealth—those with and without access difficulties were interested. Sociodemographic factors were generally not, in themselves, systematically related to telehealth interest. Therefore, older people are just as interested as their younger counterparts, after adjusting for other factors, such as confidence in using the technology. While availability of technology was quite high, this factor did not consistently relate to interest in telehealth either. The ramifications of these findings are important for policy makers, researchers, health professionals, and patients alike.

Strengths and Limitations

The main strength of this research is that it is a broad exploration of patient interest in several general forms of telehealth. The findings, therefore, are not limited to a specific intervention but highlight some of the key elements we must pay particular attention to when designing and implementing future telehealth initiatives. This is also the only study, to our knowledge, that has gathered ratings about interest in using a variety of forms of telehealth from patients that are the most likely recipients of this type of care in the future—those with chronic diseases. If telehealth is going to have an important role in effectively supporting patients with chronic diseases, then knowing what interests patients in taking up different forms of telehealth is an important first step.

The primary limitation of this study is the response rate of 44.4%, although it compares favorably with other community surveys [37]. It is possible that those who did not respond had different characteristics from those who chose to respond, which could have implications for the findings. With respect to telehealth, non-response bias by age is important given the widely held perception that older people do not like or use technology much. In this survey, responders were actually older than non-responders, and yet a fair amount of interest in telehealth was reported. There was also considerable variation in patient health and sociodemographics, as well as the other key variables of interest. These findings, nevertheless, may not be generalizable to other chronic disease patient groups. A second limitation is that the telehealth interest ratings are based on questions about hypothetical and general technologies, rather than existing or specifically named telehealth services. While this approach was directly in line with the purpose of the research—to inquire about future interest in services that *could* be delivered by existing health care providers—it is difficult to know what types of applications respondents were thinking about when they gauged their interest in telehealth delivered via social networking websites, for instance. Moreover, it is likely that there is relationship between how frequently technology is used and technology confidence [38,39], and this relationship should be controlled for in future research. Finally, the large number of variables analyzed raises the possibility of type 1 error due to multiple comparisons. Therefore, we have conservatively drawn attention only to findings that were consistent across several analyses.

Conclusions

We examined whether those with greater health needs or greater difficulties accessing traditional health care were more interested in using telehealth but found only weak evidence for this in patients with depression. Our results revealed an association between greater service delivery access difficulties (getting the right amount of care, from the right health professional, at the right time) and heightened interest in both phone-based and email/Internet-based telehealth among patients with depression. This finding aligns well with one aim of telehealth treatments for depression; namely, overcoming barriers to care. Given the results of systematic reviews that showed that telephone-administered psychotherapy [40] attrition rates are far lower than those of face-to-face care [41], this level of engagement may suggest that telehealth meets this aim to some extent. It is important to note, however, that respondents in our survey reported very few difficulties with health care access and did not report especially great health needs, except for the mental health needs of those with depression. Overall, this restriction in range may have made it difficult to detect an effect of need or a more widespread effect of access difficulties on telehealth interest.

Sociodemographic factors were found to be relatively unimportant after adjusting for attitudes towards telehealth and availability of technology, which suggests that telehealth appeals to a broader demographic than young, educated, and affluent patients. While this runs contrary to some previous literature [21,22], it might be explained by the fact that more proximal variables, like technology confidence and beliefs about telehealth, had not been included in previous research. Indeed, when similar behavioral or motivational factors are assessed, other research is consistent with our findings in demonstrating the integral role [42], or even superiority [39], of these constructs over demographic variables, albeit in terms of using the Internet alone. Furthermore, a systematic review concluded that focusing on patient factors alone, as the majority of research in this area has done, is probably not comprehensive enough to understand patient interest in using telehealth [20]. We agree with this review that future research must cut across a broader spectrum of factors, especially those at the level of human-technology interaction, the health care system, and other social or normative influences.

Technology confidence is an example of a human-technology interaction variable, and a key finding of this and other research [20] is that confidence is consistently associated with interest in using telehealth. While we asked respondents about technology confidence, it is interesting to note that other studies have used negative framing and asked about technology anxiety [38,43] or difficulty using the Internet [39]. Nonetheless, these studies also showed the equivalent association between lower technology anxiety and heightened interest in using telehealth.

There are several interesting implications of the finding that telehealth interest is most strongly associated with technology confidence and perceived advantages and disadvantages of telehealth. First, it suggests that telehealth interest is likely to increase over time as the population as a whole becomes more familiar with and comfortable using different forms of technology. This may be particularly true of social media-based telehealth [44], the newest type of technology included in the survey, and also the technology that respondents reported least confidence and interest in using. Following from this, since our results revealed that technology confidence was modality-specific, whereby confidence using one type of technology was related only to interest in using that same form of telehealth, this suggests that willingness to use telehealth is not restricted to patients who are confident using technology in general. Third, if patients were provided with adequate training and support in using telehealth equipment, they might be more interested in using telehealth. Finally, as some of telehealth's advantages are realized and other disadvantages are dispelled through telehealth use, the strength of this effect may increase. There is good reason to expect such positive experiences of telehealth, since the majority of telehealth research that asks about patient satisfaction does report fairly high levels of satisfaction [45]. Our study is no exception; high levels of satisfaction with past NHS Direct use—a form of telehealth—were reported. Positive experiences with telehealth may stimulate interest in using additional forms of telehealth, in an upward spiraling effect.

This research suggests that many people with chronic diseases are interested in using telehealth, regardless of their health status and age, and they have the technology available to them. This interest can be increased by helping people gain confidence in using technologies, highlighting the perceived advantages of telehealth, and dispelling or addressing concerns about perceived disadvantages. Based on our findings, future telehealth interventions would be best received by patients if delivered via phone or over email and static forms of Internet, rather than social media. This is because the results show that these forms of technology are readily available, patients are confident using them, and patients are most interested in telehealth delivered via these means.

Acknowledgments

We would like to thank the user groups, who volunteered through the Mental Health Research Network, for their feedback, which contributed to the final version of the questionnaire. We also thank all of the participating GP surgeries, as well as the respondents to the survey. The Healthlines Study also acknowledges the support of the National Institute for Health Research, through the Primary Care Research Network, who assisted with the recruitment of GP practices.

This article presents independent research funded by the National Institute for Health Research (NIHR) under Healthlines program (RP-PG 0108/10011). The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health.

Authors' Contributions

LE conducted the survey, undertook the analysis, and wrote the first draft of the paper. LY, AO'C, CS, and AM conceived and designed the study and obtained funding. AM supervised analysis of the survey. AG conducted background research and developed the survey questionnaire, with contributions from CT, LY, AO'C, AM, and CS. LE and Lisa Esmonde collected the data, coordinated by CT as Program Manager, under the overall guidance of CS as Chief Investigator. All authors contributed to the interpretation of the findings and critiqued the output for important intellectual content.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Subset of constructed questionnaire items.

[PDF File (Adobe PDF File), 153KB - jmir_v16i5e123_app1.pdf]

References

1. Koch S. Ubiquitous care in aging societies - a social challenge. Stud Health Technol Inform 2008;134:89-95. [Medline: 18376036]

2. Cruickshank J, Beer G, Winpenny E, Manning J. Healthcare without walls: A framework for delivering telehealth at scale. 2010 Nov. URL: http://www.2020health.org/dms/2020health/downloads/reports/2020telehealthLOW.pdf [accessed 2011-12-22] [WebCite Cache ID 647Slasr4]
3. Department of Health. Raising the profile of long term conditions care: a compendium of information. 2008. URL: http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/documents/digitalasset/dh_082067.pdf [accessed 2011-12-22] [WebCite Cache ID 647T8TbTD]
4. Wootton R. Twenty years of telemedicine in chronic disease management--an evidence synthesis. J Telemed Telecare 2012 Jun;18(4):211-220 [FREE Full text] [doi: 10.1258/jtt.2012.120219] [Medline: 22674020]
5. World Health Organization. Glossary of globalization, trade and health terms URL: http://www.who.int/trade/glossary/story021/en/ [accessed 2014-03-12] [WebCite Cache ID 6O1N7iTcT]
6. Coulton S, Coulton K, Moston V. Kent TeleHealth Evaluative Development pilot: a study into the management of people with long term conditions. 2010 Sep. URL: https://shareweb.kent.gov.uk/Documents/adult-Social-Services/professionals-and-projects/WSD/Telehealth Full Report FINAL_Layout 1.pdf [accessed 2011-12-22] [WebCite Cache ID 647UTaMKr]
7. Henderson C, Knapp M, Fernández JL, Beecham J, Hirani SP, Cartwright M, Whole System Demonstrator evaluation team. Cost effectiveness of telehealth for patients with long term conditions (Whole Systems Demonstrator telehealth questionnaire study): nested economic evaluation in a pragmatic, cluster randomised controlled trial. BMJ 2013;346:f1035 [FREE Full text] [Medline: 23520339]
8. Steventon A, Bardsley M, Billings J, Dixon J, Doll H, Hirani S, Whole System Demonstrator Evaluation Team. Effect of telehealth on use of secondary care and mortality: findings from the Whole System Demonstrator cluster randomised trial. BMJ 2012;344:e3874 [FREE Full text] [Medline: 22723612]
9. Ekeland AG, Bowes A, Flottorp S. Effectiveness of telemedicine: a systematic review of reviews. Int J Med Inform 2010 Nov;79(11):736-771. [doi: 10.1016/j.ijmedinf.2010.08.006] [Medline: 20884286]
10. Neubeck L, Redfern J, Fernandez R, Briffa T, Bauman A, Freedman SB. Telehealth interventions for the secondary prevention of coronary heart disease: a systematic review. Eur J Cardiovasc Prev Rehabil 2009 Jun;16(3):281-289. [doi: 10.1097/HJR.0b013e32832a4e7a] [Medline: 19407659]
11. Hailey D, Roine R, Ohinmaa A. The effectiveness of telemental health applications: a review. Can J Psychiatry 2008 Nov;53(11):769-778. [Medline: 19087471]
12. McManus RJ, Mant J, Bray EP, Holder R, Jones MI, Greenfield S, et al. Telemonitoring and self-management in the control of hypertension (TASMINH2): a randomised controlled trial. Lancet 2010 Jul 17;376(9736):163-172. [doi: 10.1016/S0140-6736(10)60964-6] [Medline: 20619448]
13. Kessler D, Lewis G, Kaur S, Wiles N, King M, Weich S, et al. Therapist-delivered Internet psychotherapy for depression in primary care: a randomised controlled trial. Lancet 2009 Aug 22;374(9690):628-634. [doi: 10.1016/S0140-6736(09)61257-5] [Medline: 19700005]
14. Coulter A, Ellins J. Patient-focused interventions: a review of the evidence. 2006 Aug. URL: http://www.health.org.uk/public/cms/75/76/313/526/Patient focused interventions.pdf?realName=juNCmR.pdf [accessed 2011-12-22] [WebCite Cache ID 647VLajlb]
15. Fahey T, Montgomery AA, Barnes J, Protheroe J. Quality of care for elderly residents in nursing homes and elderly people living at home: controlled observational study. BMJ 2003 Mar 15;326(7389):580 [FREE Full text] [doi: 10.1136/bmj.326.7389.580] [Medline: 12637404]
16. Paré G, Jaana M, Sicotte C. Systematic review of home telemonitoring for chronic diseases: the evidence base. J Am Med Inform Assoc 2007;14(3):269-277 [FREE Full text] [doi: 10.1197/jamia.M2270] [Medline: 17329725]
17. 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]
18. Waller R, Gilbody S. Barriers to the uptake of computerized cognitive behavioural therapy: a systematic review of the quantitative and qualitative evidence. Psychol Med 2009 May;39(5):705-712. [doi: 10.1017/S0033291708004224] [Medline: 18812006]
19. Subramanian U, Hopp F, Lowery J, Woodbridge P, Smith D. Research in home-care telemedicine: challenges in patient recruitment. Telemed J E Health 2004;10(2):155-161. [Medline: 15319045]
20. Or CK, Karsh BT. A systematic review of patient acceptance of consumer health information technology. J Am Med Inform Assoc 2009;16(4):550-560 [FREE Full text] [doi: 10.1197/jamia.M2888] [Medline: 19390112]
21. Green BB, Anderson ML, Ralston JD, Catz S, Fishman PA, Cook AJ. Patient ability and willingness to participate in a web-based intervention to improve hypertension control. J Med Internet Res 2011;13(1):e1 [FREE Full text] [doi: 10.2196/jmir.1625] [Medline: 21371993]
22. Hardiker NR, Grant MJ. Factors that influence public engagement with eHealth: A literature review. Int J Med Inform 2011 Jan;80(1):1-12. [doi: 10.1016/j.ijmedinf.2010.10.017] [Medline: 21112244]

23. British Cardiac Society, British Hypertension Society, Diabetes UK, Primary Care Cardiovascular Society, Stroke Association. JBS 2: Joint British Societies' guidelines on prevention of cardiovascular disease in clinical practice. Heart 2005 Dec;91 Suppl 5:v1-52 [FREE Full text] [doi: 10.1136/hrt.2005.079988] [Medline: 16365341]
24. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, Minhas R, Sheikh A, et al. Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. BMJ 2008 Jun 28;336(7659):1475-1482 [FREE Full text] [doi: 10.1136/bmj.39609.449676.25] [Medline: 18573856]
25. Anderson KM, Odell PM, Wilson PW, Kannel WB. Cardiovascular disease risk profiles. Am Heart J 1991 Jan;121(1 Pt 2):293-298. [Medline: 1985385]
26. Kline P. An easy guide to factor analysis. London: Routledge; 1994.
27. Office for National Statistics. 13 May Census Test 2007 (England & Wales) URL: http://www.ons.gov.uk/ons/guide-method/census/2011/the-2011-census/collecting-the-information/developing-the-questionnaires/test-questionnaire-for-england-and-wales.pdf [accessed 2011-12-22] [WebCite Cache ID 647fgLn4O]
28. Department of Health. The GP Patient Survey - Y4Q1 April-June. 2009. URL: http://www.gp-patient.co.uk/download/Questionnaires/Y4Q1 GP Patient Survey questionnaire.pdf [accessed 2011-12-22] [WebCite Cache ID 647fq3Cln]
29. Office for National Statistics. 29 April Census 2001 (England Household form H1) URL: http://www.ons.gov.uk/ons/guide-method/census/census-2001/about-census-2001/census-2001-forms/england-household-form-h1.pdf [accessed 2013-10-17] [WebCite Cache ID 6PPZnolxC]
30. The Health and Social Care Information Centre. The Health Survey for England 2007 - Household Questionnaire URL: http://www.esds.ac.uk/doc/6112/mrdoc/pdf/6112interviewingdocs.pdf [accessed 2011-12-22] [WebCite Cache ID 6PGvMceZ0]
31. 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]
32. Gulliford M, Figueroa-Munoz J, Morgan M, Hughes D, Gibson B, Beech R, et al. What does 'access to health care' mean? J Health Serv Res Policy 2002 Jul;7(3):186-188. [Medline: 12171751]
33. Rosen R, Florin D, Dixon J. Access to health care – Taking forward the findings from the scoping exercise. 2001. URL: http://www.nets.nihr.ac.uk/__data/assets/pdf_file/0003/64335/FR-08-1109-012.pdf [accessed 2011-12-22] [WebCite Cache ID 647YnDW5R]
34. Ajzen I. The theory of planned behavior. Organizational behavior and human decision processes 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
35. Gately C, Rogers A, Kirk S, McNally R. Integration of devices into long-term condition management: a synthesis of qualitative studies. Chronic Illn 2008 Jun;4(2):135-148. [doi: 10.1177/1742395308092484] [Medline: 18583451]
36. Gandek B, Ware JE, Aaronson NK, Apolone G, Bjorner JB, Brazier JE, et al. Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: results from the IQOLA Project. International Quality of Life Assessment. J Clin Epidemiol 1998 Nov;51(11):1171-1178. [Medline: 9817135]
37. 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]
38. Lai TY, Larson EL, Rockoff ML, Bakken S. User acceptance of HIV TIDES--Tailored Interventions for Management of Depressive Symptoms in persons living with HIV/AIDS. J Am Med Inform Assoc 2008;15(2):217-226 [FREE Full text] [doi: 10.1197/jamia.M2481] [Medline: 18096915]
39. Mead N, Varnam R, Rogers A, Roland M. What predicts patients' interest in the Internet as a health resource in primary care in England? J Health Serv Res Policy 2003 Jan;8(1):33-39. [doi: 10.1258/135581903360468209] [Medline: 12683432]
40. Mohr DC, Vella L, Hart S, Heckman T, Simon G. The Effect of Telephone-Administered Psychotherapy on Symptoms of Depression and Attrition: A Meta-Analysis. Clin Psychol 2008;15(3):243-253 [FREE Full text] [doi: 10.1111/j.1468-2850.2008.00134.x] [Medline: 21369344]
41. Wierzbicki M, Pekarik G. A meta-analysis of psychotherapy dropout. Professional Psychology: Research and Practice 1993;24(2):190-195. [doi: 10.1037/0735-7028.24.2.190]
42. Busselle R, Reagan J, Pinkleton B, Jackson K. Factors affecting Internet use in a saturated-access population. Telematics and Informatics 1999 Feb;16(1-2):45-58. [doi: 10.1016/S0736-5853(99)00018-0]
43. Kim J, Park HA. Development of a health information technology acceptance model using consumers' health behavior intention. J Med Internet Res 2012;14(5):e133 [FREE Full text] [doi: 10.2196/jmir.2143] [Medline: 23026508]
44. Magnezi R, Bergman YS, Grosberg D. Online activity and participation in treatment affects the perceived efficacy of social health networks among patients with chronic illness. J Med Internet Res 2014;16(1):e12 [FREE Full text] [doi: 10.2196/jmir.2630] [Medline: 24413148]
45. Williams TL, May CR, Esmail A. Limitations of patient satisfaction studies in telehealthcare: a systematic review of the literature. Telemed J E Health 2001;7(4):293-316. [doi: 10.1089/15305620152814700] [Medline: 11886667]

Abbreviations

CVD: cardiovascular disease

GP: general practitioner
MCS: Mental Component Summary
NHS: National Health Service
PCS: Physical Component Summary
SF-12v2: Short-Form Health Survey, version 2

Edited by G Eysenbach; submitted 21.01.14; peer-reviewed by B Green, D Smithard, C Bartz; comments to author 12.02.14; revised version received 14.03.14; accepted 28.04.14; published 08.05.14.

Please cite as:
Edwards L, Thomas C, Gregory A, Yardley L, O'Cathain A, Montgomery AA, Salisbury C
Are People With Chronic Diseases Interested in Using Telehealth? A Cross-Sectional Postal Survey
J Med Internet Res 2014;16(5):e123
URL: http://www.jmir.org/2014/5/e123/
doi:10.2196/jmir.3257
PMID:24811914

Original Paper

Using Online Social Media for Recruitment of Human Immunodeficiency Virus-Positive Participants: A Cross-Sectional Survey

Patrick Yuan[1], BA; Michael G Bare[1], MPH; Mallory O Johnson[1], PhD; Parya Saberi[1], PharmD, MAS

Center for AIDS Prevention Studies, Department of Medicine, University of California, San Francisco, San Francisco, CA, United States

Corresponding Author:
Parya Saberi, PharmD, MAS
Center for AIDS Prevention Studies
Department of Medicine
University of California, San Francisco
Box 0886
San Francisco, CA, 94122
United States
Phone: 1 415 597 8177
Fax: 1 415 597 9213
Email: parya.saberi@ucsf.edu

Abstract

Background: There are many challenges in recruiting and engaging participants when conducting research, especially with HIV-positive individuals. Some of these challenges include geographical barriers, insufficient time and financial resources, and perceived HIV-related stigma.

Objective: This paper describes the methodology of a recruitment approach that capitalized on existing online social media venues and other Internet resources in an attempt to overcome some of these barriers to research recruitment and retention.

Methods: From May through August 2013, a campaign approach using a combination of online social media, non-financial incentives, and Web-based survey software was implemented to advertise, recruit, and retain participants, and collect data for a survey study with a limited budget.

Results: Approximately US $5,000 was spent with a research staff designated at 20% of full-time effort, yielding 2034 survey clicks, 1404 of which met the inclusion criteria and initiated the survey, for an average cost of US $3.56 per survey initiation. A total of 1221 individuals completed the survey, yielding 86.97% retention.

Conclusions: These data indicate that online recruitment is a feasible and efficient tool that can be further enhanced by sophisticated online data collection software and the addition of non-financial incentives.

(J Med Internet Res 2014;16(5):e117) doi:10.2196/jmir.3229

KEYWORDS

HIV; AIDS; online social media; Facebook; Twitter; recruitment; Internet research; survey retention; online data collection software; non-financial incentives

Introduction

Recruiting and retaining HIV-positive research participants is a crucial yet challenging endeavor [1-4]. Numerous studies indicate several unique challenges that prevent HIV-positive individuals from participating in research, such as self-presentation bias [5], access to adequate transportation [6], low socioeconomic status [7], and mistrust of research [5]. Although offering financial incentives can mitigate some of these factors, it can also incentivize fraudulent responses and repeat participation [8], which may diminish scientific validity. Moreover, HIV-positive individuals experience high levels of perceived stigma [5,9,10], elevated rates of depression [5,7,11,12], and increased risk of substance use. When combined, these issues create a formidable barrier to participation in research designed to examine these issues and decrease the generalizability of study findings.

In addition to these HIV-specific challenges, there are major impediments and compromises associated with research in

general [13,14]. These include geographical and time constraints, labor to recruit and retain participants, limited financial resources, and logistical difficulties associated with survey administration or intervention conduction [14]. Failure to recruit and retain a sufficient number of eligible participants can be costly [15], threatens internal and external validity of the study [16], and, most importantly, deprives the scientific community of knowledge and potentially useful interventions [15,17].

Due to the near ubiquity of the Internet [18], it has become easier for people to engage in online surveys and research from the comfort and privacy of their own homes [19]. Specifically, online social media has been highly successful in capturing audiences [20]. Chief among these is Facebook, the second most popular website in the world [21]. With 198 million monthly active users and 93% of adult Internet users having a Facebook account, the average user spends approximately 1 out of every 8 minutes online or more than 11 hours per month on Facebook [22]. Twitter is another popular online social media platform, with over 49 million monthly active users [20]. These venues have already been used to successfully recruit research participants from various groups, such as mothers of advanced maternal age [23], individuals with Klinefelter's syndrome [24], those with active depression [25,26], and smokers [27,28].

Therefore, if perceived stigma, cost of transportation, and physical distance are potential deterrents for HIV-positive individuals to participate in research, and time, geographical, and resource constraints are among the general difficulties of research, then Internet-based resources may be a logical solution to overcome many of these barriers. Thus far, the literature has indicated that online social media is ripe for researchers to use as a tool for recruitment [29,30]. By accessing online venues where potentially qualified participants are already spending time, it is possible to bridge the psychosocial and physical divides between research participants and researchers to achieve a successful outcome. Therefore, to capitalize on these existing resources, we used a novel combination of online social media along with non-financial incentives and contemporary data collection software for participant recruitment, retention, and data collection.

Methods

Study Design and Data Collection

We conducted a cross-sectional survey using online social media to recruit HIV-positive individuals. The purpose of the parent study was to examine the barriers to and facilitators of adherence to antiretroviral therapy and to describe the use of mobile telephones and other technologies to improve adherence among HIV-positive individuals living in the United States. Here, we describe the methods used for participant recruitment using online social media.

We implemented a campaign approach where participants were recruited through online social media such as Facebook, Twitter, and other social media platforms, such as LinkedIn, Craigslist, and Tumblr, from May 7 to August 31, 2013. The online survey was programmed using Qualtrics Research Suite (Qualtrics, Provo, UT), an online survey tool that allows researchers to build, distribute, and analyze online surveys in real time. The survey contained 112 items on demographics, HIV clinical outcomes, and use of technology. The University of California, San Francisco (UCSF) Committee on Human Research approved this study in April 2013.

Participant Recruitment

A schedule was devised by the research assistant to regularly update and maintain these social media sites for a total of 8 hours per week (20% of full-time effort).

Facebook

Overview

We used 4 methods to recruit participants on Facebook: paid advertisements (ads), fan page, personal messages, and postings in groups.

Facebook Ads

A Facebook ad was created that ran separately and concurrently with the rest of the research recruitment campaign to further broaden exposure to the target population. With the assistance of a Facebook representative, a daily average budget of US $40 and a total budget of US $5,000 were established to cover the 4-month recruitment period. The average maximum bid per cost per click (CPC) was set at US $3.50, which was above the Facebook suggested bid range of US $0.52-$1.09, to outcompete other advertisements targeting the same groups. These advertisements were targeted at users with self-reported interests or "liked" topics in any of the following categories: HIV/AIDS; lesbian, gay, bisexual, transgender, queer, questioning, intersex, and asexual (LGBTQQIA); men who have sex with men (MSM); HIV co-morbidities such as tuberculosis and Hepatitis C; and unprotected sex.

A wide variety of advertisement pictures were chosen to cater to the target demographic. Pictures of inanimate or abstract objects, such as the AIDS red ribbon, were used as well as pictures of people representing various age ranges, genders, and sexual identities. Since banner advertisements targeting MSM of color have been shown to increase the click-through rate [31], we altered the racial/ethnic composition of the images. All of these advertisements used the banner title "Living with HIV?", accompanied by a brief survey description (Multimedia Appendix 1). We obtained the rights to use stock photographs from various Internet sites.

We initiated the Facebook ad campaign with 3 advertisements but increased the number of advertisements to 5 in order to cater 2 advertisements specifically toward women and youth (18-29 years old). This decision was in response to the Facebook Ad Reports, which showed low levels of participation by women and youth. We varied ad pictures on a bi-weekly basis based on the number of Facebook ad clicks.

Facebook Fan Page

A Facebook fan page was created with the intention of recruiting participants and raising general awareness of the study. This was achieved by generating interesting and relevant posts for people living with HIV, which included news articles, study

announcements, survey dissemination requests, pictures, videos, and HIV-related resources. Posted news articles related to HIV, medication adherence, and the use of technology as it pertained to HIV and medication adherence. On average, the research assistant updated the fan page every 3 to 4 days and posted about 2 new posts per day. A fan page description under the "About" section of the Facebook page was created to inform potential participants of the study and provide a link to the survey. A brief description of the study investigators and the organization was also included to build rapport and credibility with the audience. The research assistant interacted with other Facebook fan pages with similar purpose or interest by "liking" these organizations' pages, which were found using keyword searches for HIV, AIDS, MSM, and health care.

Facebook Messages

The research assistant sent personalized Facebook messages on behalf of the study to community leaders, major HIV/AIDS organizations, and other related organizations, especially in states with low survey response rates. Survey response rates were estimated by examining Facebook Ad Reports. These messages contained information about the study, a link to the survey, and a request to repost the study link on the individual or organization's corresponding Facebook page.

Facebook Groups

The research assistant also joined HIV-related Facebook groups to advertise for recruitment. These groups maintain Facebook pages based on real-life interests to facilitate discussions of relevant topics. Once requests to join these groups were granted, posts were made in the common thread requesting that interested participants click on the survey link.

Twitter

A Twitter account was created for the purpose of recruiting participants for the study and to continue building an online network to advertise for the survey. Through the study's Twitter account (ie, "feed"), the research assistant followed any organization affiliated with topics such as HIV, AIDS, MSM, or global health care. Individuals were also followed if they were candid about their HIV seropositive status or being MSM, an HIV advocate, or a health educator.

Tweets were sent directly to both individuals and organizations as a request to re-tweet the survey link. Hashtags (#) related to the Twitter account were used to keep track of the number of times each message was tweeted and a shorter version of the study's URL (ie, "tiny URL") was generated to fit within the 140-character limit of tweets. The research assistant tweeted relevant content related to HIV/AIDS or interacted with other Twitter feeds by re-tweeting interesting and applicable topics on a weekly basis.

Other Social Media Platforms

LinkedIn

LinkedIn groups that included content related to HIV were specifically chosen to advertise to potential study participants. Postings regarding the study were made in groups that were created for people living with HIV, HIV advocates, community workers, and other researchers. Additionally, a request was made with each posting for group members to share the survey link with other potentially interested groups or individuals.

Craigslist

On a weekly basis, the research assistant posted advertisements containing the survey link and basic information about the study on nationwide Craigslist groups pertaining to MSM and health. Elements from the Touro 12-Step Process [32] were utilized, including guidelines on creating a thread title, generating an active and ongoing discussion, and being courteous.

Tumblr

A Tumblr blog was created and maintained with regular updates that mirrored the Facebook posts. Other blogs with content related to HIV were "liked" in order to build an audience and stimulate recruitment.

Subjects and Research Engagement

Individuals who clicked on the study Web link were directed to the Qualtrics survey where they were required to consent online before being screened for eligibility. The survey was used to screen individuals based on the following inclusion criteria: age 18 years or older, HIV-positive serostatus, and currently living in the United States. We excluded individuals who had already taken the survey by limiting the number of survey attempts to 1 per Internet Protocol (IP) address and by asking participants if they had already taken the survey. Qualtrics was programmed to end the survey if the participant did not qualify. We ensured anonymity by not collecting any identifiers or personal health information and storing responses on encrypted and password-protected servers at UCSF.

Given the anonymous nature of the Web-based survey and lack of identifiers to verify repeat and fraudulent responses, monetary incentives were not offered in order to de-incentivize multiple survey attempts. However, in order to keep participants engaged, we motivated participants by inserting 5 medically interesting facts strategically throughout the survey titled "Fun Facts", which were accompanied by visually stimulating and relevant pictures. Examples of these fun facts include: "Did you know that we exercise at least 30 muscles when we smile?" and "Did you know that your brain uses as much power as a 10-watt light bulb?" Additionally, to spark interest, we asked participants before beginning the survey if they knew of a natural substance that can be potentially effective against HIV. Upon completion of the survey, we embedded a link to a video about how bee venom is being studied by the Washington University School of Medicine as a potential drug to treat HIV [33].

We consulted the UCSF Center for AIDS Prevention Studies (CAPS) Community Advisory Board (CAB), which is comprised of stakeholders from Bay Area agencies and communities with the mission of channeling community input into the CAPS research agenda, providing advice on study methods, and improving recruitment methods. Based on the recommendations of the CAPS CAB, we modified the Facebook ad pictures to exclusively use photographs of people, the Facebook ad text to be more direct and motivating, and incorporated Tumblr as an additional online social networking venue for recruitment.

Analysis

Three variables were created to estimate the relative value of each of the 3 main methods of recruitment: (1) "Facebook Page Interactions" to estimate user activity on the Facebook fan page, (2) "Twitter Interactions" to evaluate the overall activity on the Twitter feed, and (3) "Facebook Ad Clicks" to assess the overall success of the Facebook ad. The number of likes, comments, and shares on the Facebook page per recruitment day were extracted from Facebook Page Insights and summed to serve as an objective measure of Facebook Page Interactions. Data collected from Twitonomy (a Twitter analytics tool) included the number of retweets and mentions per recruitment day and were used to compute Twitter Interactions. Information on the number of clicks per day for the Facebook ad was obtained from Facebook Ad Reports. Using descriptive data from Qualtrics, we collected information on the number of survey clicks per day. We used Pearson's pairwise correlation to examine the relationship between the number of survey clicks per day and Facebook Page Interactions, Twitter Interactions, and Facebook Ad clicks per day. Correlation values of 0 to .3 were deemed weak correlation, .3 to .7 indicated moderate correlation, and >.7 were considered strong correlation. The significance level of each correlation coefficient was also estimated. A *P* value of <.05 was deemed statistically significant.

As part of the survey, participants were also asked about their method of recruitment. These methods were categorized under one of the following categories: Facebook, Twitter, word-of-mouth, email from a listserv, and other (including LinkedIn, Craigslist, Tumblr, or other). We used 2-way frequency tables to examine the distribution of the method of recruitment based on participants' age (categorized as 18-29.9, 30-39.9, 40-49.9, 50-59.9, and ≥60 years), race/ethnicity (White/Caucasian, African-American/Black, Latino, and other), male sex at birth, and sexual orientation (homosexual/gay, heterosexual, bisexual, and other). A chi-square test was used to assess the null hypothesis that the categorical variables are independent.

Last, to estimate research retention in the absence of monetary incentives, we compared the total number of participants who responded to the first survey question with the total number of respondents who answered the last survey question. All statistical analyses were conducted using Stata, version 13.1 (StataCorp, College Station, TX).

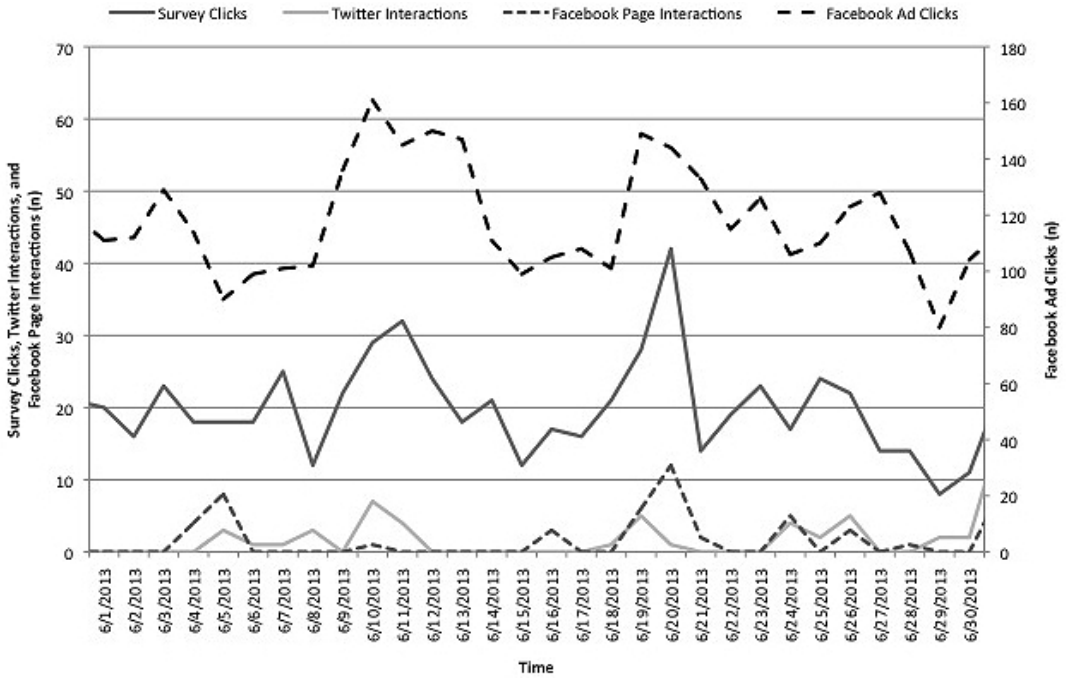
Results

Total Respondents

A total of 2034 individuals clicked on the survey Web link, averaging 18 clicks per day. Of those who clicked, 1977 people consented to take the survey, 1404 met the inclusion criteria and initiated the first question of the survey, and 1221 responded to the last question of the survey (86.97% retention). A total of 43 respondents answered affirmatively when asked if they had taken the survey before and were therefore excluded from the survey.

The study received a total of 10,006 Facebook Ad Clicks, 278 Facebook Page Interactions, and 161 Twitter Interactions during the recruitment period. Although each of these factors contributed to the total number of survey clicks per day, only the number of Facebook Ad Clicks was moderately correlated with the number of survey clicks per day based on the Pearson pairwise correlation (.52, *P*<.001). The Facebook Page Interactions and Twitter Interactions were weakly correlated with Facebook Ad Clicks (*r*=.17, *P*=.06; *r*=.18, *P*=.06, respectively). Figure 1 illustrates the number of survey clicks relative to the number of Facebook Ad clicks, Twitter Interactions, and Facebook Page Interactions for the month of June 2013. This serves to visually demonstrate the correlation between each of the 3 major methods of recruitment and the total number of survey clicks per day.

Figure 1. Number of Survey Clicks, Twitter Interactions, Facebook Interactions, and Facebook Ad Clicks in June 2013.



Number of Survey Clicks, Twitter Interactions, Facebook Interactions, and Facebook Ad Clicks in June 2013

On average, participants spent 31 minutes to complete the survey. Table 1 encapsulates key demographics of participants by self-reported method of recruitment. According to these data, participants were primarily 40-49.9 years of age, identified as White/Caucasian, male at birth, and homosexual. Facebook was the most common method of recruitment across all reported demographics. With increase in age, participants were more likely to report being recruited through non-Facebook methods. Similarly, heterosexual individuals were more likely to report non-Facebook methods of recruitment compared to bisexual or homosexual participants. Email from a listserv was the second most commonly reported method for those aged 60 years or older and African Americans. Racial/ethnic categories were independent of self-reported method of recruitment.

A total of US $5,021 was spent on the Facebook ad (mean CPC = $0.64), which targeted a total of 14 million Facebook users and accumulated 1404 initiated surveys. Therefore, the estimated mean cost per survey initiation was approximately US $3.56. The reach of the Facebook ads included participants from 48 states plus the District of Columbia but did not include participants from either Alaska or Wyoming. From the Facebook fan page, the research assistant "liked" 490 other fan pages while 53 other organizations and individual Facebook users "liked" the study's fan page. A total of 141 Facebook messages were sent to community leaders, organizations, and individuals requesting to have the survey link reposted.

Over the course of the recruitment period, the research assistant tweeted 572 times, received 199 Twitter followers, and followed 1092 Twitter feeds. In addition, the research assistant contacted a total of 16 LinkedIn groups, added 3 Tumblr sites, and posted on 3 Craigslist forums.

Table 1. Key participant demographics and self-reported method of recruitment.

Characteristic	Total, n	Facebook	Email from a Listserv	Word-of-mouth	Twitter	Other[a]	*P* value[b]
Age, years, n (%)							<.001
18-29.9	246	211 (85.8)	7 (2.9)	15 (6.1)	2 (0.8)	11 (4.5)	
30-39.9	274	233 (85.0)	11 (4.0)	21 (6.7)	5 (1.8)	4 (1.5)	
40-49.9	508	422 (83.1)	29 (5.7)	36 (7.1)	6 (1.2)	15 (3.0)	
50-59.9	474	381 (80.4)	37 (7.8)	33 (7.0)	5 (1.1)	18 (3.8)	
≥60	402	256 (63.7)	55 (13.7)	31 (7.7)	23 (5.7)	37 (9.2)	
Total[c]	1904	1503 (78.94)	139 (7.30)	136 (7.14)	41 (2.15)	85 (4.46)	
Racial identification, n (%)							.311
White/Caucasian	995	802 (80.6)	67 (6.7)	81 (8.1)	11 (1.1)	34 (3.4)	
Black/African American	126	97 (77.0)	14 (11.1)	9 (7.1)	2 (1.6)	4 (3.2)	
Hispanic/Latino	181	157 (86.7)	8 (4.4)	9 (5.0)	2 (1.1)	5 (2.8)	
Other	102	90 (88.2)	4 (3.9)	3 (3.0)	2 (2.0)	3 (2.9)	
Total[c]	1404	1146 (81.62)	93 (6.62)	102 (7.26)	17 (1.21)	46 (3.28)	
Sex at birth, n (%)							<.001
Male	1321	1102 (83.42)	76 (5.75)	86 (6.51)	14 (1.06)	43 (3.26)	
Female	77	39 (50.7)	16 (20.8)	16 (20.8)	3 (3.9)	3 (3.9)	
Total[c]	1398	1141 (81.62)	92 (6.58)	102 (7.30)	17 (1.22)	46 (3.29)	
Sexual orientation, n (%)							<.001
Homosexual/Gay	1208	1007 (83.36)	73 (6.04)	80 (6.62)	11 (0.91)	37 (3.06)	
Heterosexual	92	54 (58.7)	13 (14.1)	14 (15.2)	5 (5.4)	6 (6.5)	
Bisexual/Other	91	75 (82.4)	6 (6.6)	8 (8.8)	1 (1.1)	1 (1.1)	
Total[c]	1391	1136 (81.67)	92 (6.61)	102 (7.33)	17 (1.22)	44 (3.16)	

[a]Includes LinkedIn, Craigslist, and Tumblr.

[b]Based on chi-square test.

[c]Differences in number of responses for each characteristic are due to missing data; participants were not "forced" to respond to each demographic question.

Discussion

Principal Findings

Our study results indicate that the unique integration of online social media recruitment, Web-based survey, and non-financial incentives is an efficient and streamlined strategy for survey research with HIV-positive individuals. The restricted resources of US $5,000 in a 4-month recruitment timeframe with 20% research assistant's time resulted in recruiting 1404 qualified participants in 48 states and the District of Columbia, averaging 18 survey responses per day, making this an efficient solution for participant recruitment. This is similar to other studies not focused on an HIV-positive population that have demonstrated the efficiency of Facebook ads [24,27]. Furthermore, we were able to access a difficult-to-reach population that may experience stigma and other barriers to research participation.

Even though Facebook has been used to effectively recruit HIV-positive individuals [34], our study is novel in that we utilized a campaign approach—using several different online social media platforms at once to benefit from various approaches and to maximize diversity. Compared to other methods of online recruitment, the Facebook ad was by far the most successful and least time consuming. The number of survey clicks per day and Facebook ad clicks was moderately correlated and highly statistically significant, illustrating the importance of a paid Facebook advertisement. We believe that setting a high CPC bid and increasing the number of interest groups as recommended by the Facebook representative, as well as using ad images of women and ethnic minorities [32] and more direct and motivating ad text as proposed by the CAPS CAB were associated with improved ad performance.

Although we observed a weak correlation between the number of survey clicks per day and both the Facebook Page and Twitter Interactions, these platforms were essential in establishing and maintaining a rapport with the study population and community. Such relationships were cultivated by liking and following relevant pages and groups of underrepresented populations, posting comments on walls and reposting relevant material, and sending personal messages. These techniques were particularly helpful given that the level of compliance with our requests was contingent upon the relationship and rapport between the researcher and the person or institution being contacted. For example, mass-produced messages and random posts on other Facebook pages or Twitter feeds were not as useful as personalized messages sent directly to a community leader or an HIV advocate with an established relationship. This highlights the social phenomenon known in economics as reciprocity, which correlates one's likelihood of compliance with their familiarity with the one making the request [35]. Therefore, future researchers should not rely exclusively on online social media and underestimate the powerful nature of existing social networks to assist with recruitment efforts.

Retaining participants for Internet-based studies has been known to be difficult [36], particularly among African American and Latino MSM [31,37]. In response, our study catered ads for these demographics and implemented a novel incentive structure by adding interesting facts throughout the survey to motivate participants. We used this approach as opposed to one with pecuniary incentives to minimize duplicate and false responses. Given that the average response time for survey completion was 31 minutes and 86.97% of those who initiated the first question of the survey completed the survey, we believe this is a tactical approach that can be used in future online research with anonymous participants.

Since there are more Facebook users than any other social media site, it was not surprising that most participants identified Facebook as their primary method of recruitment. Moreover, the paid Facebook ad and the selection of targeted groups likely contributed to a higher percentage of Facebook recruits. However, it is unknown why older age and non-homosexual orientations were more likely to be associated with non-Facebook methods of recruitment. Future research should further examine possible reasons for these patterns. As such, the use of online social networks other than Facebook for recruitment remains a strong area of interest as venues for future research [38]. It is likely that sites such as LinkedIn, Craigslist, and Tumblr would be more useful for research within professional networks, other interest groups, or longitudinal research.

Limitations

The self-reported demographics elucidated in Table 1 are not representative of the population of those living with HIV in the United States [39]. Responses revealed a higher proportion of individuals who identify as homosexual but fewer African Americans and Latinos. This may be attributed to HIV-specific reasons such as stigma, the necessity of having access to the Internet to participate in the survey, the need to be using social media to view recruitment advertisements, and the requirement for English proficiency to take the survey. Such drawbacks may be generalized to other study populations [24,40] and are important factors when considering the use of online social media for recruitment.

Another limitation of our study is the possibility of duplicate responses. A study using online social media also found this issue to be problematic, especially when participants were provided a financial incentive without the use of restrictive software [8]. However, attrition rates for online surveys in the HIV-positive MSM population are high [41], urging researchers to provide some form of incentive. With awareness of this issue, we implemented safeguards to protect the integrity of the survey data by introducing non-financial incentives, automatically disqualifying participants who used duplicate IP addresses, and asking participants whether they had already participated in the survey. This inquiry was a necessary precaution owing to the fact that 43 respondents admitted to taking the survey more than once. Future researchers might consider investigating the extent to which repeat online survey attempts are denied due to duplicate IP addresses using comparable online survey instruments and non-financial incentives. Even though we believe the likelihood of duplicate responses is low, it is not possible for us to determine if users circumvented these measures by accessing the survey through multiple devices or by providing inaccurate responses.

In terms of Facebook ad pictures, future research may be directed toward studying the effectiveness of different ad images by measuring the number of ad clicks and survey responses, which was a limitation of this study. Another limitation was that we were unable to verify the HIV serostatus of participants. However, given the lack of pecuniary incentives, it is unlikely that individuals would misrepresent themselves as being HIV-positive. Last, given that email from a listserv and directly contacting individuals using Facebook messages were valuable tools, it would have been beneficial for us to have asked participants about the specific listserv or the particular Facebook route (ie, Facebook ads vs personal message) in order to distinguish between various efforts.

Conclusions

Online social media is an indispensable tool for recruiting participants because it has the potential to be cost-effective and efficient. Social media has the unique ability to transcend barriers to study recruitment such as physical distance, transportation, and limited time and financial resources. It also affords users the added benefits of privacy and anonymity, which may ameliorate the effects of perceived stigma in difficult to reach populations. Moreover, non-financial incentives in the form of trivia are a viable alternative to monetary incentives that yields high rates of retention while minimizing the chances of duplicate and fraudulent responses. Facilitating this process using online data collection software can further reduce cost and time associated with data collection. In sum, as Internet-based research tools become more practical, researchers must embrace these methods not only to maximize efficiency but also to enhance the scientific validity and generalizability of their findings. As such, online social media, alternative types of incentives, and Web-based survey software are well poised to become staple methods of recruitment and engagement in research.

Acknowledgments

The researchers would like to thank Mr Nikolai Caswell for his help in editing this manuscript. This project was supported by NIH award numbers K23MH097649, K24MH087220, and P30AI027763. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Facebook ad example.

[PNG File, 66KB - jmir_v16i5e117_app1.png]

References

1. Anastasi JK, Capili B, Kim GH, Chung A. Clinical trial recruitment and retention of a vulnerable population: HIV patients with chronic diarrhea. Gastroenterol Nurs 2005;28(6):463-468. [Medline: 16418582]
2. McQuiston C, Uribe L. Latino recruitment and retention strategies: community-based HIV prevention. J Immigr Health 2001 Apr;3(2):97-105. [doi: 10.1023/A:1009565900783] [Medline: 16228793]
3. Calsyn RJ, Klinkenberg WD, Morse GA, Miller J, Cruthis R, HIV/AIDS Treatment Adherence, Health Outcomes Cost Study Group. Recruitment, engagement, and retention of people living with HIV and co-occurring mental health and substance use disorders. AIDS Care 2004;16 Suppl 1:S56-S70. [Medline: 15736822]
4. Dodds S, Blakley T, Lizzotte JM, Friedman LB, Shaw K, Martinez J, et al. Retention, adherence, and compliance: special needs of HIV-infected adolescent girls and young women. J Adolesc Health 2003 Aug;33(2 Suppl):39-45. [Medline: 12888286]
5. DiClemente RJ, Ruiz MS, Sales JM. Barriers to adolescents' participation in HIV biomedical prevention research. J Acquir Immune Defic Syndr 2010 Jul;54 Suppl 1:S12-S17 [FREE Full text] [doi: 10.1097/QAI.0b013e3181e1e2c0] [Medline: 20571418]
6. Saberi P, Yuan P, John M, Sheon N, Johnson MO. A pilot study to engage and counsel HIV-positive African American youth via telehealth technology. AIDS Patient Care STDS 2013 Sep;27(9):529-532. [doi: 10.1089/apc.2013.0185] [Medline: 23991691]
7. Traeger L, O'Cleirigh C, Skeer MR, Mayer KH, Safren SA. Risk factors for missed HIV primary care visits among men who have sex with men. J Behav Med 2012 Oct;35(5):548-556 [FREE Full text] [doi: 10.1007/s10865-011-9383-z] [Medline: 22068878]
8. Quach S, Pereira JA, Russell ML, Wormsbecker AE, Ramsay H, Crowe L, et al. The good, bad, and ugly of online recruitment of parents for health-related focus groups: lessons learned. J Med Internet Res 2013;15(11):e250 [FREE Full text] [doi: 10.2196/jmir.2829] [Medline: 24231040]

9. Mahajan AP, Sayles JN, Patel VA, Remien RH, Sawires SR, Ortiz DJ, et al. Stigma in the HIV/AIDS epidemic: a review of the literature and recommendations for the way forward. AIDS 2008 Aug;22 Suppl 2:S67-S79 [FREE Full text] [doi: 10.1097/01.aids.0000327438.13291.62] [Medline: 18641472]
10. Sayles JN, Wong MD, Kinsler JJ, Martins D, Cunningham WE. The association of stigma with self-reported access to medical care and antiretroviral therapy adherence in persons living with HIV/AIDS. J Gen Intern Med 2009 Oct;24(10):1101-1108 [FREE Full text] [doi: 10.1007/s11606-009-1068-8] [Medline: 19653047]
11. Bing EG, Burnam MA, Longshore D, Fleishman JA, Sherbourne CD, London AS, et al. Psychiatric disorders and drug use among human immunodeficiency virus-infected adults in the United States. Arch Gen Psychiatry 2001 Aug;58(8):721-728. [Medline: 11483137]
12. Valente SM. Depression and HIV disease. J Assoc Nurses AIDS Care 2003;14(2):41-51. [Medline: 12698765]
13. Maxine P, Victor D, Lakshika T. Challenges in recruitment of research participants. Advances in Psychiatric Treatment 2003 May;9:229-238.
14. Hulley SB. Designing Clinical Research. Philadelphia, PA: Lippincott Williams & Wilkins; Feb 2013, July.
15. Nasser N, Grady D, Balke CW. Commentary: Improving participant recruitment in clinical and translational research. Acad Med 2011 Nov;86(11):1334-1335 [FREE Full text] [doi: 10.1097/ACM.0b013e3182302831] [Medline: 22030634]
16. Johnson MO, Dilworth SE, Neilands TB, Chesney MA, Rotheram-Borus MJ, Remien RH, NIMH HLP team. Predictors of attrition among high risk HIV-infected participants enrolled in a multi-site prevention trial. AIDS Behav 2008 Nov;12(6):974-977 [FREE Full text] [doi: 10.1007/s10461-007-9356-y] [Medline: 18202908]
17. Hogan JW, Roy J, Korkontzelou C. Handling drop-out in longitudinal studies. Stat Med 2004 May 15;23(9):1455-1497. [doi: 10.1002/sim.1728] [Medline: 15116353]
18. Pew Internet & American Life Project. Internet adoption. 2013. URL: http://www.pewinternet.org/Static-Pages/Trend-Data-(Adults)/Internet-Adoption.aspx [accessed 2013-12-27] [WebCite Cache ID 6MBbyYNks]
19. Zhang D, Bi P, Lv F, Tang H, Zhang J, Hiller JE. Internet use and risk behaviours: an online survey of visitors to three gay websites in China. Sex Transm Infect 2007 Dec;83(7):571-576 [FREE Full text] [doi: 10.1136/sti.2007.026138] [Medline: 17971376]
20. Business Insider. Twitter has a surprisingly small number of US users. 2013. URL: http://www.businessinsider.com/twitter-has-a-surprisingly-small-number-of-us-users-2013-10 [accessed 2014-01-06] [WebCite Cache ID 6MPzXGaDx]
21. Alexa. Top sites in United States. 2010. URL: http://www.alexa.com/topsites/countries/US [accessed 2014-01-06] [WebCite Cache ID 6MPzZxD3v]
22. Hubspot. 12 Essential Facebook Stats. 2011 URL: http://blog.hubspot.com/blog/tabid/6307/bid/14715/12-Essential-Facebook-Stats-Data.aspx [accessed 2014-01-06] [WebCite Cache ID 6MPzbPjjw]
23. O'Connor A, Jackson L, Goldsmith L, Skirton H. Can I get a retweet please? Health research recruitment and the Twittersphere. J Adv Nurs 2014 Mar;70(3):599-609. [doi: 10.1111/jan.12222] [Medline: 23909740]
24. Close S, Smaldone A, Fennoy I, Reame N, Grey M. Using information technology and social networking for recruitment of research participants: experience from an exploratory study of pediatric Klinefelter syndrome. J Med Internet Res 2013;15(3):e48 [FREE Full text] [doi: 10.2196/jmir.2286] [Medline: 23512442]
25. Morgan AJ, Jorm AF, Mackinnon AJ. Internet-based recruitment to a depression prevention intervention: lessons from the Mood Memos study. J Med Internet Res 2013;15(2):e31 [FREE Full text] [doi: 10.2196/jmir.2262] [Medline: 23403043]
26. 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]
27. 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]
28. Sadasivam RS, Volz EM, Kinney RL, Rao SR, Houston TK. Share2Quit: Web-based peer-driven referrals for smoking cessation. JMIR Res Protoc 2013;2(2):e37 [FREE Full text] [doi: 10.2196/resprot.2786] [Medline: 24067329]
29. Ryan GS. Online social networks for patient involvement and recruitment in clinical research. Nurse Res 2013;21(1):35-39. [doi: 10.7748/nr2013.09.21.1.35.e302] [Medline: 24004430]
30. 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]
31. Sullivan PS, Khosropour CM, Luisi N, Amsden M, Coggia T, Wingood GM, et al. Bias in online recruitment and retention of racial and ethnic minority men who have sex with men. J Med Internet Res 2011;13(2):e38 [FREE Full text] [doi: 10.2196/jmir.1797] [Medline: 21571632]
32. Ip EJ, Barnett MJ, Tenerowicz MJ, Perry PJ. The Touro 12-Step: a systematic guide to optimizing survey research with online discussion boards. J Med Internet Res 2010;12(2):e16 [FREE Full text] [doi: 10.2196/jmir.1314] [Medline: 20507843]
33. Hood JL, Jallouk AP, Campbell N, Ratner L, Wickline SA. Cytolytic nanoparticles attenuate HIV-1 infectivity. Antivir Ther 2013;18(1):95-103. [doi: 10.3851/IMP2346] [Medline: 22954649]

34. Mitchell JW, Petroll AE. Patterns of HIV and sexually transmitted infection testing among men who have sex with men couples in the United States. Sex Transm Dis 2012 Nov;39(11):871-876 [FREE Full text] [doi: 10.1097/OLQ.0b013e3182649135] [Medline: 23060078]
35. Fehr E, Gächter S. Fairness and retaliation: the economics of reciprocity. Journal of Economic Perspectives 2000 Aug;14(3):159-182. [doi: 10.1257/jep.14.3.159]
36. Koo M, Skinner H. Challenges of internet recruitment: a case study with disappointing results. J Med Internet Res 2005;7(1):e6 [FREE Full text] [doi: 10.2196/jmir.7.1.e6] [Medline: 15829478]
37. Khosropour CM, Sullivan PS. Predictors of retention in an online follow-up study of men who have sex with men. J Med Internet Res 2011;13(3):e47 [FREE Full text] [doi: 10.2196/jmir.1717] [Medline: 21745792]
38. Van de Belt TH, Berben SA, Samsom M, Engelen LJ, Schoonhoven L. Use of social media by Western European hospitals: longitudinal study. J Med Internet Res 2012;14(3):e61 [FREE Full text] [doi: 10.2196/jmir.1992] [Medline: 22549016]
39. Centers for Disease Control and Prevention. HIV/AIDS. 2013. URL: http://www.cdc.gov/hiv/statistics/basics/ataglance.html [accessed 2014-02-22] [WebCite Cache ID 6NaIaSNRu]
40. Eysenbach G, Wyatt J. Using the Internet for surveys and health research. J Med Internet Res 2002;4(2):E13 [FREE Full text] [doi: 10.2196/jmir.4.2.e13] [Medline: 12554560]
41. Chiasson MA, Parsons JT, Tesoriero JM, Carballo-Dieguez A, Hirshfield S, Remien RH. HIV behavioral research online. J Urban Health 2006 Jan;83(1):73-85 [FREE Full text] [doi: 10.1007/s11524-005-9008-3] [Medline: 16736356]

Abbreviations

CAPS CAB: Center for AIDS Prevention Studies Community Advisory Board
CPC: cost per click
IP: Internet Protocol
MSM: men who have sex with men
UCSF: University of California, San Francisco

Edited by G Eysenbach; submitted 06.01.14; peer-reviewed by J Kwong, M Koo; comments to author 31.01.14; revised version received 11.03.14; accepted 30.03.14; published 01.05.14.

Please cite as:
Yuan P, Bare MG, Johnson MO, Saberi P
Using Online Social Media for Recruitment of Human Immunodeficiency Virus-Positive Participants: A Cross-Sectional Survey
J Med Internet Res 2014;16(5):e117
URL: http://www.jmir.org/2014/5/e117/
doi:10.2196/jmir.3229
PMID:24784982

Original Paper

Use of the Internet as a Health Information Resource Among French Young Adults: Results From a Nationally Representative Survey

François Beck[1,2], PhD; Jean-Baptiste Richard[1], MSc; Viet Nguyen-Thanh[1], MSc; Ilaria Montagni[3,4]; Isabelle Parizot[4,5], PhD; Emilie Renahy[6], PhD

[1]Institut National de Prévention et d'Education pour la Santé (INPES), Paris, France

[2]Cermes3 - Equipe Cesames (Centre de recherche Médecine, Sciences, Santé, Santé mentale, Société), Université Paris Descartes, Sorbonne Paris Cité /CNRS UMR 8211/Inserm U988/EHESS, 45, rue des Saints-Pères. 75270 Paris Cedex 06, Paris, France

[3]Section of Psychiatry, Department of Public Health and Community Medicine, University of Verona, Verona, Italy

[4]Research Team on the Social Determinants of Health and Healthcare, UMRS 707, INSERM, Paris, France

[5]CNRS, Centre Maurice Halbwachs, ERIS, Paris, France

[6]Centre for Research on Inner City Health, Li Ka Shing Knowledge Institute, St. Michael's Hospital, Toronto, ON, Canada

Corresponding Author:
François Beck, PhD
Institut National de Prévention et d'Education pour la Santé (INPES)
42, Bld de la Libération
Saint Denis Cedex
Paris, 93203
France
Phone: 33 01 49 33 23
Fax: 33 01 49 33 23
Email: francois.beck@inpes.sante.fr

Abstract

Background: The Internet is one of the main resources of health information especially for young adults, but website content is not always trustworthy or validated. Little is known about this specific population and the importance of online health searches for use and impact. It is fundamental to assess behaviors and attitudes of young people looking for online health-related information and their level of trust in such information.

Objective: The objective is to describe the characteristics of Internet users aged 15-30 years who use the Web as a health information resource and their trust in it, and to define the context and the effect of such use on French young adults' behavior in relation to their medical consultations.

Methods: We used the French Health Barometer 2010, a nationally representative survey of 27,653 individuals that investigates population health behaviors and concerns. Multivariate logistic regressions were performed using a subsample of 1052 young adults aged 15-30 years to estimate associations between demographics, socioeconomic, and health status and (1) the use of the Internet to search for health information, and (2) its impact on health behaviors and the physician-patient relationship.

Results: In 2010, 48.5% (474/977) of Web users aged 15-30 years used the Internet for health purposes. Those who did not use the Internet for health purposes reported being informed enough by other sources (75.0%, 377/503), stated they preferred seeing a doctor (74.1%, 373/503) or did not trust the information on the Internet (67.2%, 338/503). However, approximately 80% (371/474) of young online health seekers considered the information found online reliable. Women (P<.001) and people with higher sociocultural positions (OR 0.5, 95% CI 0.3-0.9 and OR 0.4, 95% CI 0.2-0.7 for employees and manual workers, respectively, vs individuals with executive or manager positions) were more likely to use the Internet for health purposes. For a subsample of women only, online health seeking was more likely among those having a child (OR 1.8, 95% CI 1.1-2.7) and experiencing psychological distress (OR 2.0, 95% CI 1.0-4.0). Finally, for online health seekers aged 15-30 years, one-third (33.3%, 157/474) reported they changed their health behaviors (eg, frequency of medical consultations, way of taking care of one's own health) because of their online searches. Different factors were associated with different outcomes of change, but psychological distress, poor quality of life, and low income were the most common.

Conclusions: The Internet is a useful tool to spread health information and prevention campaigns, especially to target young adults. Young adults trust online information and consider the Internet as a valid source of health advice. Health agencies should ensure the improvement of online health information quality and the creation of health-related websites and programs dedicated to young adults.

(J Med Internet Res 2014;16(5):e128) doi:10.2196/jmir.2934

KEYWORDS

health communication; information dissemination; access to information; Internet; trust; young adults

Introduction

Background

The use of the Internet to look for advice or health information has been a growing resource since the 1990s [1]. Health prevention programs can benefit from the Internet especially when dedicated to or designed for young adults who represent the vast majority of Web users [2]. High-quality health information can be provided through websites, forums, blogs, and social networks, which have been some of the most popular channels for health promotion among young people in the past 10 years [2,3]. In France, 3 of 4 people have access to the Internet, and Internet use is higher in young people compared to other adults: 99% of people aged 12-17 years use the Internet, and this proportion falls to 22% for those aged 70 years and older [4]. Some French websites exclusively address health issues concerning either the general population or adolescents in particular. However, given the pace with which informal websites and blogs are created [5], young Web users do not exclusively use official websites whose content is trustworthy and certified by experts and quality labels [6,7]. For this reason, and given the amount of health information available on the Internet, it is fundamental to investigate behaviors and attitudes of young adults searching for health-related information on the Web (young online health seekers or aged 15-30 years health seekers). Therefore, we deemed it important to describe the profile of young online health seekers together with the context and consequences of their searches.

The Web offers a large amount of health-related information and benefits from different interactive formats. However, the disparate quality of available information [8-11] might reinforce social disparities among Web users [12]. This heterogeneity is also linked to the perception of reliability and credibility Web users have regarding the information found on the Internet [13]. The French National Authority for Health (Haute Autorité de Santé; HAS) pilots the certification procedure of health-related websites by using the Health On the Net (HON) Code [6]. However, this initiative does not provide a complete evaluation of all available information. According to the Pew Research Center's Internet & American Life Project conducted in the United States in 2009, 3% of online health seekers stated they had health problems after having followed medical advice or information found on the Web [14]. Ambiguity regarding the quality of health information on the Web affects and worries some adult Web users [15]. Therefore, it is essential to understand what young people think about the credibility of online health information.

Furthermore, the effect of these searches on the way young people take care of their own health and well-being is still unknown. American studies report that, in the general population, the Internet is used to get additional information and/or advice about one's own health [14], namely when facing a diagnosis and/or having to choose a treatment [15,16]. Sometimes the use of the Internet can postpone and even replace medical consultation and treatment [14]. Although teenagers do not usually make decisions about medical care autonomously, it is still relevant to assess the impact these online searches have on young adults' health behaviors.

In France, several studies have focused on health information-seeking on the Internet [4,17-19]. However, to the best of our knowledge, no nationally representative sample has provided in-depth analysis of the behaviors and perceptions of young online health seekers by focusing on gender, income, and socioeconomic and health status.

Objectives

The aims of this article are: (1) to provide information about the prevalence of Internet use for health-related purposes in France among young adults and define the sociodemographic, socioeconomic, and health-related profile of users, (2) to investigate the context and the impact of the information found on health-related behaviors, and (3) to assess the level of trust young adults have in the information found on the Internet.

Methods

Survey Methodology

Data were extracted from the National French Health Barometer survey conducted in 2010 by the French Institute for Prevention and Health Education (INPES) in consultation with the French Ministry of Health [20]. This survey was designed to measure the evolution of key indicators regarding health-related behaviors, attitudes, and opinions in the general population. Using a computer-assisted telephone interviewing (CATI) system, 27,653 people were interviewed from October 22, 2009 to July 3, 2010. Interviewers from a private survey firm were trained by the INPES to administrate this health-related survey.

We used a 2-stage random sampling design: (1) selection of households using random digit dialing covering all metropolitan French regions, and (2) random selection of one member of the household, using the method proposed by Kish [21]. Because of the increasing rate of households that have abandoned their landline telephones for cell phones, a cell-only sample was added (12% of the sample to keep the same rate as in 2010 in France). The cell-only sample was created independently from

the landline sample by using the prefix (2 digits) assigned by the National Telecom Authority to each mobile phone provider. The remaining digits of the phone numbers (8 digits) were then randomly generated. If the respondent had a landline phone in his/her household, he/she was excluded from the cell-only sample. This development was essential to improve the coverage rate [22] because of the number of dwellings with a landline phone (87% in 2010 vs 96% in 1998). Thus, approximately 99% of the population was covered [4]. Details of the survey methodology have been published previously elsewhere [23,24].

If a household or individual refused to participate or could not be reached, they were not replaced in the study. Thus, specific efforts were made to successfully reach households and increase the response rate: a formal request to participate explaining the goals of the study was sent by mail before the first call (addresses were located from the landline phone numbers when available); unsuccessful calls were repeated after 30 and 90 minutes, on different days, and at different times to a maximum of 40 attempts for each generated phone number. Individuals who refused to participate were contacted a second time by specially trained interviewers. The overall refusal rate was 39%. All collected data were anonymous and self-reported. The mean duration of an interview was approximately 32 minutes for landline phones and 34 minutes for mobile phones.

This population-based survey procedure was approved by the French data protection authority (Commission Nationale de l'Informatique et des Libertés; CNIL), an independent administrative body that operates in accordance with the national data protection legislation, amended in 2004 specifically to protect citizens' identities and privacy and ensure access to their own personal data.

From the initial nationally representative sample of 27,653 people aged 15-85 years, a random sample of 1052 young adults aged 15-30 years answered a set of specific questions on their use of the Internet as an information tool for health-related issues. In this paper, we will analyze this subsample (referred to as young adults or aged 15-30 years interchangeably).

Data were weighted by the number of telephone lines and eligible persons in the household. They were also adjusted to represent the French population structure (2008 census) according to age, gender, educational level, region of residence, and level of urbanization.

Independent Variables

Sociodemographic Characteristics

Sociodemographic characteristics included the following: age group (15-19 years, 20-25 years, and 26-30 years), gender, socio-occupational status (categorized as manual workers, employees, intermediate occupations, executive and manager positions, and other), and income by consumption unit (adjusting for the household size and divided into quintiles). For those not working at the time of the interview, we used the head of household's socio-occupational status.

Health Status

Respondents were questioned about their health status and if they had children or were expecting a child. The Duke Index, a validated tool containing 17 items that assesses general health status [25], was used to measure respondents' health and well-being (score ranging from 0 to 100, later analyzed in tertiles). Psychological distress was measured using the 5-item Mental Health scale (MH-5; using a validated cut-off of 55), a specific section from the Short-Form 36 (SF-36) questionnaire, which is a validated, multipurpose, short-form health questionnaire with 36 questions [26]. Presence of a chronic disease was assessed by a self-reported answer (yes/no); if they answered yes, the disease had to be specified. Moreover, a variable named "fear of illness" was created as a score, analyzed in quartiles, summing answers to 10 questions concerning fear (not at all, a little, quite a few, a lot) of specific diseases or events (eg, traffic accidents, alcohol diseases, cancer, Alzheimer disease). Level of information on health issues was measured with a series of 13 questions (eg, Do you feel you are well informed about alcohol/tobacco/cancer...?). A 4-item scale (very well/well/bad/very badly informed) was used and the total score was analyzed using quartiles.

Trust in Internet Information

Survey respondents were asked about the credibility and trustworthiness of health-related information obtained on the Internet. Responses were categorized as reliable, somewhat reliable, not really reliable, not reliable at all, and "do not know."

Dependent Variables

Use of the Internet as a Source of Health Information

Survey respondents were asked whether they had ever used the Internet to search for information and advice about health and the frequency of their search(es) (eg, "During the past 12 months have you used the Internet to look for information or advice about health?" and "If so, how many times per week, month, or year?"). We also asked about the themes of their searches to a randomized subsample of 139 online health seekers. Answers to the latter question were grouped into 5 categories: general health and illnesses, medical news and treatments, mother and child health, health behaviors, and occasional diseases.

Individuals who never looked for health information on the Internet were asked if this was because they had enough information through other resources, they were not interested in getting health information, they were more confident in seeing a doctor for health-related questions, they were not confident with the information provided on the Internet, or they never thought about using the Internet to search for health-related information.

The Effect of Using the Internet on the Doctor-Patient Relationship

We subsequently asked the subsample of online health seekers if the information and advice found on the Internet had changed the way they take care of their health. In addition, they were asked if the use of the Internet led them to visit their doctor more often, less often, or as they did before using the Internet for health purposes. The context of the search was investigated by analyzing if respondents had often (compared with rarely or never) used the Internet for health purposes in the following situations: instead of seeing a doctor, before seeing a doctor,

after having seen a doctor, and without link to any medical consultation.

Statistical Analysis

Bivariate chi-square tests were performed on weighted proportions considering the following thresholds: .001, .01, and .05. Five multivariate logistic regression models were used to investigate whether risk factors (listed previously) were associated with (1) the use of the Internet for health purposes, and (2) the context and consequences of online health searches: Internet search instead of seeing a doctor, change in taking care of one's own health, or having seen a physician less/more frequently. We estimated adjusted odds ratio (adjusted OR) and 95% confidence intervals (95% CI) based on the Wald test. The analyses were performed using R-3.0.1 software.

Results

Sample Characteristics

Data used in this analysis included 1052 individuals aged 15-30 years, of which 50.48% (531/1052) were men and 49.52% (521/1052) were women (Table 1). The mean age was 22.6 years (SD 0.18).

Table 1. Participant characteristics of landline and cell-only samples.

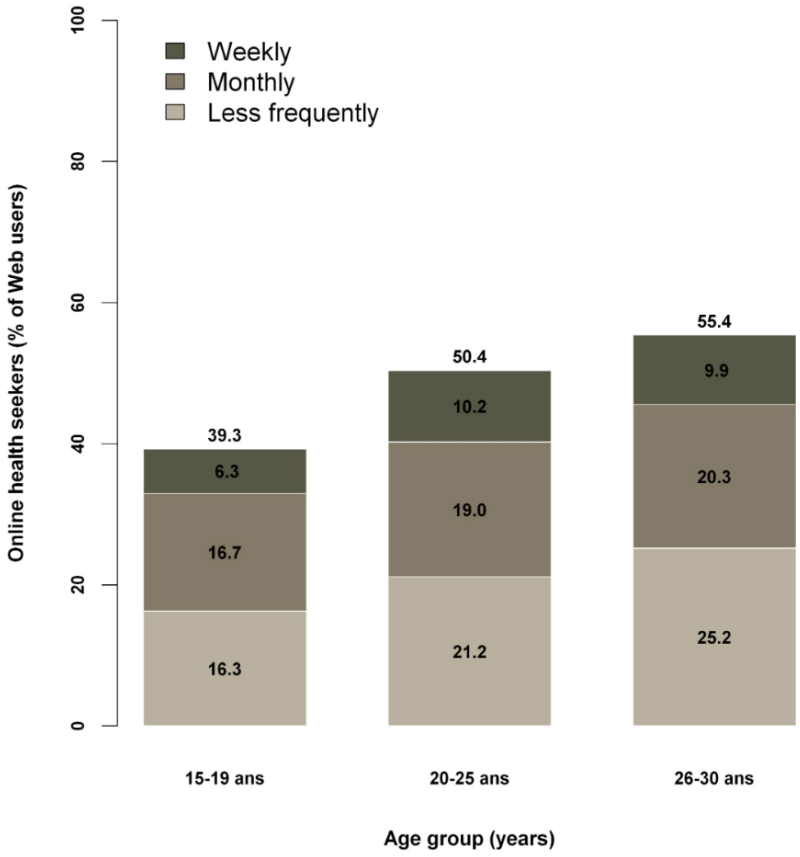
Characteristics	Total (N=1052)	Landline sample (n=733)	Cell-only sample (n=319)
Gender, n (%)			
Men	531 (50.48)	369 (50.3)	162 (50.8)
Women	521 (49.52)	364 (49.7)	157 (49.2)
Age (years), n (%)			
15-19	322 (30.61)	289 (39.4)	33 (10.4)
20-25	391 (37.17)	235 (32.1)	156 (48.9)
26-30	339 (32.22)	209 (28.6)	130 (40.7)

Use of the Internet as a Source of Health Information

Overview

Almost all (977/1052, 92.87%) of our sample of young adults was comprised of Web users, and this proportion decreased slightly as age of respondents increased (from 96.2% to 90.1% for the 15-19 years and 26-30 years age groups, respectively) (see Figure 1).

Figure 1. Web users and online health seekers by age group.



Nonhealth Seekers

Among 977 Web users, 503 (51.5%) had never used the Internet to look for health information and advice during the past 12 months: 75.0% (377/503) explained being adequately informed by other sources, 74.1% (373/503) preferred seeing a doctor, and 67.2% (338/503) did not trust the information found on the Internet. Although there was no statistically significant difference by age, the youngest group (15-19 years) seemed more likely than others to be adequately informed through other sources and to have a distrust in information found on the Internet (Table 2).

Table 2. Reasons for not using the Internet for health information among Web users by age.

Reasons	Age group (years), n (%)				*P*
	15-30 (n=503)	15-19 (n=188)	20-25 (n=179)	26-30 (n=136)	
Adequately informed by other means and resources	377 (75.0)	148 (78.7)	128 (71.5)	101 (74.3)	.52
Not interested in this type of information	203 (40.4)	75 (39.9)	76 (42.5)	52 (38.2)	.80
More confident in seeing a doctor for this kind of information	373 (74.1)	138 (73.4)	131 (73.1)	104 (76.4)	.69
Distrust in the information provided by the Internet	338 (67.2)	135 (71.8)	114 (63.7)	89 (65.4)	.38
Do not know	244 (48.5)	102 (54.2)	78 (43.6)	64 (47.1)	.21

Health Seekers

Half of the Web users (474/977, 48.5%) used the Internet during the past 12 months to look for either information or advice on health: 8.9% (87/977) every week, 18.7% (183/977) every month, and 20.9% (204/977) less frequently (Figure 1). In summary, 45% of young adults used the Internet for health purposes.

Among Internet users, the use of the Internet for health purposes significantly increased with age: 39.3%, 50.4%, and 55.4% for the 15-19 years, 20-25 years, and 26-30 years age groups, respectively (*P*=.002). After adjusting for socioeconomic and health status, and all variables presented in Table 3, the logistic models showed that the likelihood of using the Internet for health purposes was higher among women compared to men (57.2% vs 39.7%; OR 1.8, 95% CI 1.3-2.3; *P*<.001; result not shown). Employees and manual workers were less likely than executives and managers to search for online health information and they actually surfed the Internet (whatever the reason) less frequently (among 15-30 year age group, 96.4% of executives

and managers vs 92.6% of employees and 86.7% of manual workers). Finally, women with psychological distress (OR 2.0, 95% CI 1.0-4.0; *P*=.046), as well as pregnant women or women with at least 1 child (OR 1.8, 95% CI 1.1-2.7; *P*=.01) used the Internet for health purposes more frequently than others (Table 3).

Table 3. Factors associated with the use of the Internet for health purposes among Web users aged 15-30 years by gender.

Variables	Men (n=487)				Women (n=490)			
	n (%)	Adjusted OR	95% CI	*P*	n (%)	Adjusted OR	95% CI	*P*
Socio-occupational status				.04				.10
Other	34 (54.0)	0.6	0.3, 1.4	.28	26 (47.6)	0.5	0.2, 1.4	.20
Executives and managers	80 (54.5)	1			57 (64.8)	1		
Intermediate occupations	129 (41.2)	0.7	0.4, 1.2	.15	163 (65.8)	0.9	0.4, 1.7	.64
Employees	92 (33.6)	0.5	0.3, 0.9	.03	172 (51.4)	0.5	0.2, 0.9	.02
Manual workers	152 (33.5)	0.4	0.2, 0.7	<.001	72 (54.7)	0.5	0.2, 1.0	.04
Quality of life (Duke Index)				.89				.03
Third tertile (poor)	70 (40.6)	1			161 (62.3)	1		
Second tertile (medium)	163 (37.9)	0.8	0.5, 1.5	.57	172 (61.3)	1.0	0.6, 1.6	.96
First tertile (good)	254 (40.6)	1.0	0.6, 1.8	.93	157 (47.3)	0.6	0.4, 1.1	.09
Chronic disease				.23				.40
Yes	40 (50.0)	1			46 (50.1)	1		
No	447 (38.8)	0.7	0.3, 1.3	.22	444 (57.8)	1.3	0.7, 2.5	.38
Psychological distress				.51				.17
No	467 (39.4)	1			426 (56.0)	1		
Yes	20 (47.9)	1.6	0.6, 4.4	.34	63 (66.4)	2.0	1.0, 4.0	.046
Having a child/being pregnant				.12				.002
No	432 (38.3)	1			356 (52.9)	1		
Yes	55 (50.7)	1.7	0.9, 3.0	.09	134 (69.7)	1.8	1.1, 2.7	.01

[a]Logistic regression models were adjusted on all shown variables.

Themes of Online Health Searches

For the searched themes, those aged 15-30 years primarily looked for information on general health or specific diseases, especially flu or influenza (44.6%, 62/139). Themes searched by older people (31-85 years) concerned health behaviors, children's health, and parental health. Women appeared to be particularly concerned with themes concerning children's health and parental health (26.8%, 22/82) (Table 4).

Table 4. Health-related search themes according to gender among individuals aged 15-30 years.

Health topics	All, n (%) (n=139)[a]	Men, n (%) (n=57)	Women, n (%) (n=82)	*P*
General health and illnesses	62 (44.6)	27 (47.4)	35 (42.7)	.52
Children's health and parental health	29 (20.9)	6 (10.5)	22 (26.8)	.06
Specific health problems	27 (19.4)	14 (24.6)	14 (17.1)	.31
Health behaviors	27 (19.4)	12 (21.1)	14 (17.1)	.59
Medical news/care	17 (12.2)	7 (12.3)	10 (12.2)	.88

[a]A randomized subsample of 139 online health seekers who were asked to specify the content of their searches.

Trust in Online Health Information

Approximately 80% (78.2%, 371/474) of online health seekers aged 15-30 years trusted the information they found on the Internet, even if 61.4% of them (291/474) qualified the information only as "somewhat" reliable, without significant differences according to age and gender. However, the opinion on the credibility of such information was associated with the way respondents took care of their own health. Among online health seekers, those who thought the information was not really reliable were less inclined to change the way they take care of their own health than those who found online information to be

reliable (12.1% vs 39.2%, *P*<.001). Moreover people who found the information not really reliable did not decrease the frequency of their medical consultations (0.8% vs 8.1% for those who found the information reliable, *P*<.001).

The Effect of Online Health Searches on the Doctor-Patient Relationship

Table 5 illustrates the overall impact of Internet use on the young adults' medical consultations. Almost 3 of 10 online health seekers aged 15-30 years reported having often used the Internet as a source of health information instead of seeing a doctor (29.9%, 142/474) or before seeing a doctor (28.7%, 136/474). By contrast, 16.7% (79/474) used the Internet after having seen a doctor, which significantly varied by age group: the 26-30 years group looked for information on the Internet after having seen a doctor significantly more often (22.4%, *P*=.03) than the 15-19 years (13.1%) and the 20-25 years (13.7%) groups.

Moreover, a total of 33.1% (157/474) of the 15-30 years group of health seekers stated they changed their way of taking care of their health. For 11.4% (54/474) of young online health seekers, the information found on the Internet in the past 12 months led them to see a doctor more often (4.9%, 23/474) or less often (6.5%, 31/474) than usual: the 20-25 years group tended to see their doctors less frequently (9.9%, 18/182) than the 15-19 years (4.1%, 5/122) and the 26-30 years (4.7%, 8/170) groups.

Finally, although 26.6% (126/474) looked for online health information without having had any kind of medical consultation, 33.1% (157/474) reported they modified the way they take care of their health based on the information they found on the Internet (no further significant difference by age group).

Table 5. Impact of online health searches among online health seekers by age group.

Impact of online health searches	Age group (years), n (%)				*P*
	15-30 (n=474)	15-19 (n=122)	20-25 (n=182)	26-30 (n=170)	
Use the Internet for health purposes often" or very often...					
...instead of seeing a doctor	142 (29.9)	31 (25.4)	57 (31.3)	54 (31.8)	.42
...before seeing a doctor	136 (28.7)	31 (25.4)	54 (29.7)	51 (30.0)	.69
...after having seen a doctor	79 (16.7)	16 (13.1)	25 (13.7)	38 (22.4)	.03
...not in relation to a medical consultation	126 (26.6)	28 (22.9)	48 (26.4)	50 (29.4)	.39
Use the Internet for health purposes has changed the way of taking care of one's health	157 (33.1)	43 (35.2)	66 (36.3)	48 (28.2)	.35
Use the Internet for health purposes has made medical consultations...					
...more frequent	23 (4.9)	7 (5.7)	10 (5.5)	6 (3.5)	.25
...less frequent	31 (6.5)	5 (4.1)	18 (9.9)	8 (4.7)	
...as often as usual	420 (88.6)	110 (90.2)	154 (84.6)	156 (91.8)	

Table 6 presents estimates of multivariate logistic regressions of 4 different outcomes assessing the perceived impact of online health searches. Young adults reporting the lowest level of economic resources were more likely to see their physician less frequently (OR 2.7, 95% CI 1.4-5.5; *P*=.004). Those reporting poor quality of life according to the Duke scale (third tertile) were more likely to search often for health information on the Internet instead of seeing a doctor than those reporting good quality of life (OR 2.3, 95% CI 1.4-3.7; *P*<.001). They were also more likely to change the way they take care of their own health (OR 1.8, 95% CI 1.1-2.9; *P*=.009) and were more likely to see their physician more frequently (OR 2.7, 95% CI 1.0-7.4; *P*=.048). Finally, people with psychological distress, fearing illnesses (OR 3.3, 95% CI 1.0-10.5; *P*=.04), and those less informed about diseases (OR 3.2, 95% CI 1.1-9.0; *P*=.02) tended to increase the frequency of their consultations because of their online health searches.

Table 6. Factors associated with the context and consequences of online health searches, with odds ratios adjusted (adj OR) on all shown variables (N=474).

Factor	n	Internet search instead of seeing a doctor			Change in taking care of one's own health			Has seen the physician less frequently			Has seen the physician more frequently		
		Adj OR	95% CI	*P*	Adj OR	95% CI	*P*	Adj OR	95% CI	*P*	Adj OR	95% CI	*P*
Gender													
Male	193	1			1			1			1		
Female	281	0.7	0.5-1.1	.11	0.6	0.4-0.8	.003	0.5	0.3-1.1	.06	1.9	0.8-4.7	.15
Low income (1st quintile)													
No	363	1			1			1			1		
Yes	111	1.0	0.71.6	.89	1.4	0.9-2.1	.21	2.7	1.4-5.5	.004	1.7	0.8-3.7	.23
Quality of life (Duke Index)													
Third tertile(poor)	130	2.3	1.4-3.7	<.001	1.8	1.1-2.9	.009	1.2	0.5-2.7	.69	2.7	1.0-7.4	.048
Second tertile (medium)	162	1.3	0.8-2.0	.23	1.1	0.7-1.7	.79	0.7	0.3-1.7	.48	1.9	0.7-5.3	.24
First tertile (good)	182	1			1			1			1		
Chronic disease													
Yes	44	1			1			1			1		
No	430	0.8	0.4-1.4	.41	0.7	0.4-1.2	.14	0.9	0.3-2.6	.81	0.7	0.2-2.0	.42
Psychological distress													
No psychological distress	419	1			1			1			1		
Psychological distress	55	2.2	1.3-3.7	.004	1.0	0.6-1.8	.77	1.1	0.4-3.1	.73	3.2	1.4-7.4	.003
Fear of illness													
First quartile (less afraid)	115	1			1			1			1		
Second quartile	129	1.4	0.8-2.3	.22	1.0	0.6-1.7	.94	1.3	0.5-3.0	.53	1.4	0.4-5.0	.64
Third quartile	129	1.3	0.8-2.1	.39	1.6	1.0-2.7	.04	0.5	0.2-1.4	.19	1.9	0.6-6.6	.26
Fourth quartile (more afraid)	102	1.2	0.7-2.0	.64	1.7	1.0-2.9	.06	0.9	0.3-2.3	.65	3.3	1.0-10.5	.047
Level of information on health issues													
First quartile (well informed)	96	1			1			1			1		
Second quartile	129	1.2	0.7-2.1	.60	1.1	0.6-1.8	.80	1.4	0.4-4.3	.56	0.8	0.2-2.9	.75
Third quartile	122	1.6	0.9-2.8	.12	0.6	0.3-1.0	.048	1.2	0.4-3.9	.88	0.7	0.2-2.6	.65
Fourth quartile (poorly informed)	127	1.9	1.1-3.3	.03	1.0	0.6-1.7	.96	2.4	0.8-6.8	.16	3.2	1.1-9.0	.02

Discussion

Characteristics of Online Health Seekers

According to the national Health Barometer 2010 survey data based on a random and representative sample of the French population, 45% of young adults aged 15-30 years used the Internet in the past 12 months to seek health information. This result is in-line with a study from the French National Institution of Statistics and Economical Studies (INSEE) in 2010 examining the same question, among others, but over a period of 3 months [17]. At the international level, this proportion is lower than in other countries. A survey performed in 7 European countries (Norway, Denmark, Germany, Greece, Poland, Portugal, and Latvia) in 2005 found that, on average, 63% of individuals aged 18-29 years were online health seekers [27]. The replication of this study in 2007 showed that this behavior was growing in all age groups [28]. Another survey carried out in 2010 in Italy showed that 60% of young Italian males and 65% of young Italian females (between 18-29 years) used the Internet for health-related purposes [29]. In the United States, the Pew Research Center's Internet & American Life Project showed that in 2012, 72% of people aged 18-29 years were online health seekers [30]. Considering the 18-29 years group in our own study, the proportion of online health seekers only increased to 48%. However, this apparent lower proportion of young online health seekers in France may be because of an underestimation in our study as a result of formulation issues (see Limitations). Moreover, the aforementioned international surveys exclusively focused on online health information seeking and do not aim at being representative of the general population, which is the case with the Health Barometer 2010.

Online health seekers aged 15-30 years were more likely to be women than men and have a position as an executive or manager rather than being employees or manual workers. The literature confirms that gender, occupation, and socioeconomic status are the main factors discriminating the Internet use for health concerns in the general population [22-26]. Those factors have also been found in very different contexts and countries (eg, Saudi Arabia, Brazil, or Japan) [31-33].

Regarding the gender effect, this result is not specific to the Internet use for seeking answers to health questions because women, in general, tend to be more interested in health than men [34]. However, it is worth noting that women's use of the Internet for health advice or information seeking has empowered them and changed their relationship with health care providers [35]. Moreover, other results showed that women were more likely than men to seek help for someone else [36].

The association with socioeconomic status is also consistent with the fact that the Internet, as any technological innovation, tends to primarily benefit the wealthier and/or more educated [37-39], and in this case reinforce the inverse information law (as an extension of the inverse care law as defined in 1971 [40]): the availability and use of health information (and the ability to use it properly) tends to vary inversely with the need of the population. Understanding the sociodemographic and socioeconomic profiles of online health seekers could, therefore, help improve the quality of online information and tools (eg, by adjusting the level of literacy required) and produce age- or gender-specific online supports.

Although our sample is relatively homogeneous regarding age, we saw that the proportion of Web users slightly decreased with increasing age, whereas the proportion of online health seekers increased. This was previously found in the general population [3,41-44] and could probably be related to the ambiguity of the age effect [45]. On the one hand, younger people, namely adolescents, have more access to the Internet and have better and more Web-related skills [46-48]. On the other hand, they are less concerned with health problems, the latter increasing with age and impacting those generations who are a priori less at ease with using the Internet [49,50]. Our results suggest that national and regional health agencies could develop health promotion campaigns and programs targeting young adults to bridge the gap between their low level of knowledge regarding health issues and the increasing prevalence of lifestyle diseases [51].

Some health-related factors were also associated with the use of the Internet to search for health-related information. Our data did not show significant associations with general health status (measured in our study through chronic disease or quality of life) as often shown in the general population [27,45,52,53]. However, 2 specific conditions were found to be associated with online health searches. Psychological distress appeared related to searching for health information on the Internet. This could be explained by the fact that a specific condition, rather than a perceived general health status, increases the interest and the need to search for specific information or treatment. Moreover, anxiety itself could lead these people to look for further health information or to verify information after a medical consultation. Furthermore, the confidentiality of the Internet could represent an advantage for its use as a tool to obtain information on stigmatizing issues, such as many mental health illnesses. This could explain, at least partially, why the use of the Internet for health purposes is positively associated to poor mental health but not to physical health.

The last factor associated with the use of the Internet for health-information seeking is having or expecting a child, especially among women. Again, this might be related to the fact that women, more than men, still tend to take care of the family's health. The interest in health information dealing with parenthood is clear when we look at the most frequent themes. Questions about mother's and child's health are indeed the most frequently mentioned topics among the young Web users (21%) after general health and illnesses (45%). These findings are consistent with those in other countries in Europe, such as Italy [54]. This noticeable interest in parenthood probably represents an interesting starting point for health promotion providers and policy makers in France. The creation of specific websites on this topic could meet the needs of parents and provide them with validated information.

Context and Impact of Online Health Searches

With regard to the context of health-information seeking on the Internet, three-quarters of online health seekers reported having made their Internet searches in conjunction with a medical consultation, either before (eg, to see if a consultation is needed or to get prepared to an eventual treatment) or after (eg, to get additional information or seek for alternative treatments). More interestingly, approximately 3 in 10 young adults reported having looked for health information on the Internet instead of seeing a doctor. This behavior could fit with a search cost model using the Internet as a resource to reduce health care and information search costs [55]. By finding reassuring information about specific and precise questions on the Internet, young adults could save the money and time of any medical consultation.

Moreover, one-third of the 15-30 years group of online health seekers reported having modified the way they take care of their health after their Internet searches. It is possible that the changes in question reflect an increased distancing from health professionals, which may lead young adults to follow advice against public health rules (eg, purchasing medicines on the Internet or trusting uncontrolled therapies). Conversely, these findings could be positive if people have been trained and influenced by trustworthy online information campaigns or prevention programs. In both cases, these results confirm the idea that the Internet potentially supports the dissemination of health information with an impact on young adults' health, as well as the importance of promoting labels to guarantee the reliability of the information provided on commercial websites, or to train users to read in a critical manner.

Trust in Online Health Information

Although two-thirds of young people did not look for health information on the Internet because of their distrust in this kind of information, the majority of young online health seekers (approximately 80%) trusted the information they found on the

Internet. This could be worrisome because the quality and validity of health information on the Internet varies a lot in France, as it does in other countries [8-12]. For instance, a 2009 study conducted in the United Kingdom showed that only 4 of 10 websites provided correct information regarding pediatric issues [11]. However, our statistical analyses showed that the opinion on the credibility of online information is linked to the repercussions Internet searches have on the way people take care of their own health. In fact, those who trusted the less health-related information are also those who stated less change in their health and medical behavior because of their Internet searches. In any case, it is important to underline that we do not know to what extent the level of trust in the information found online is related to the actual validity of the information.

Finally, if young adults feel comfortable using the Internet, they may have difficulty judging the quality of health-related information or they may not be aware of quality labels. Therefore, it is fundamental to help young people to find and use the most valid online health information. Several strategies can be developed to reach this goal. On the one hand, institutional websites need to be created—or the promotion of labels on other websites—where health information is clearly thought through, well planned, referenced, and safely managed [56]. This process is already in place in the United States and Australia [2,3,43], where young Web users represent a large proportion of online health seekers. In France, the INPES also developed many validated information resources on the Internet and social networks dedicated to young people during the last 10 years. This agency also promoted its reliable online resources through other media (eg, television, schools). In a complementary manner, it seems useful to offer health-related educational programs and e-learning activities to young adults.

Another strategy to ensure young adults get valid information about health issues is to target their main way of using the Internet, namely social networks. There is a growing use of social networks for health promotion purposes and the literature shows that those interventions are effective in some fields, such as sexual health promotion [57-59]. That is also something the INPES has tried to develop in recent years. Social networks could also be established as a place for physicians and health professionals to help their patients wade through online information and make recommendations on reliable sources. It is then necessary to develop the monitoring, validation, and labeling of new tools created by health professionals and experts. Professional organizations could attempt to build digital resources for young people and work with them in a collaborative manner, as the nature of Web 2.0 suggests.

Limitations

Analyses were based on a large sample representative of the French population. The methodology of the survey has been validated and interviews were conducted by trained interviewers. However, several limitations deserve attention in the interpretation of our findings. The response rate was 61%, which is satisfactory compared with other health surveys in France, but lower than the rates obtained in other epidemiologic surveys, such as the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) [60]. However, selection bias cannot be ruled out and some populations, especially the most deprived ones (ie, homeless people), are likely to be underrepresented, although some were interviewed as a result of the sample based on mobile phone numbers.

Among the limits of our approach, it is also necessary to underline that our data do not allow for the distinction among Web sites, blogs, and social networks. People who use these tools do not necessarily all have the same approach; therefore, they might not have the same profile (eg, socioeconomic status and health behaviors). It is possible that a part of the young adults who looked for information about health behavior answered no to the question “have you used the Internet to look for health information or advice.” Tobacco smoking, sexual behaviors, drug consumption, or sleep habits may indeed not be perceived as health behaviors by young people who might not perceive the health consequences of their behaviors, particularly those that will occur in the long term.

Conclusions

Our study shows that in France in 2010, almost all individuals aged 15-30 years were Web users, and approximately half of them used the Internet to look for health information for themselves, their relatives, or nobody in particular. These results justify the increasing effort over the past several years by health promotion stakeholders in designing specific e-tools, such as the development by agencies or labeled stakeholders of websites or Facebook pages dedicated to adolescents and young adults, of online publishing of video events (eg, INPES manga [61] aimed at preventing the initiation of smoking), or the development of smartphone apps (eg, Alcoholometer app to estimate daily alcohol consumption).

To conclude, the Internet is assuming an increasingly important role in its young users’ lives and is increasingly becoming one of the major health information mediums in many countries. This explains why effective health interventions for young people should not avoid online tools. Given the results of this study, France is expected to maintain enhancing the number and quality of health-related websites especially addressed to individuals aged 15-30 years. It is incumbent to find more creative ways to inform young people about health and health care in ways that reflect their own style and culture.

Conflicts of Interest

None declared.

References

1. Baker L, Wagner TH, Singer S, Bundorf MK. Use of the Internet and e-mail for health care information: results from a national survey. JAMA 2003 May 14;289(18):2400-2406. [doi: 10.1001/jama.289.18.2400] [Medline: 12746364]
2. Burns JM, Morey C. Technology and young people's mental and well being. In: Challenges in Adolescent Health: An Australian Perspective. Victoria: Internation Academic Press; 2008:61-71.
3. Burns JM, Davenport TA, Durkin LA, Luscombe GM, Hickie IB. The internet as a setting for mental health service utilisation by young people. Med J Aust 2010 Jun 7;192(11 Suppl):S22-S26. [Medline: 20528703]
4. Bigot R, Croutte P. La diffusion des technologies de l'information et de la communication dans la société française. Paris: Credoc; 2011:244.
5. Wantland DJ, Portillo CJ, Holzemer WL, Slaughter R, McGhee E. The effectiveness of Web-based vs. non-Web-based interventions: a meta-analysis of behavioral change outcomes. J Med Internet Res 2004 Nov 10;6(4):e40 [FREE Full text] [doi: 10.2196/jmir.6.4.e40] [Medline: 15631964]
6. Nabarette H, Romaneix F, Boyer C, Darmoni SJ, Rémy PL, Caniard E. Certification des sites dédiés à la santé en France. Dossier thématique Presse Med 2009;38:1476-1483.
7. Health on the Internet. HONcode URL: http://www.hon.ch/HONcode/ [accessed 2014-04-17] [WebCite Cache ID 6Otzm6Ddu]
8. Lawrentschuk N, Abouassaly R, Hackett N, Groll R, Fleshner NE. Health information quality on the internet in urological oncology: a multilingual longitudinal evaluation. Urology 2009 Nov;74(5):1058-1063. [doi: 10.1016/j.urology.2009.05.091] [Medline: 19758687]
9. Ream E, Blows E, Scanlon K, Richardson A. An investigation of the quality of breast cancer information provided on the internet by voluntary organisations in Great Britain. Patient Educ Couns 2009 Jul;76(1):10-15. [doi: 10.1016/j.pec.2008.11.019] [Medline: 19179036]
10. Reavley NJ, Jorm AF. The quality of mental disorder information websites: a review. Patient Educ Couns 2011 Nov;85(2):e16-e25. [doi: 10.1016/j.pec.2010.10.015] [Medline: 21087837]
11. Scullard P, Peacock C, Davies P. Googling children's health: reliability of medical advice on the internet. Arch Dis Child 2010 Aug;95(8):580-582. [doi: 10.1136/adc.2009.168856] [Medline: 20371593]
12. Cline RJ, Haynes KM. Consumer health information seeking on the Internet: the state of the art. Health Educ Res 2001 Dec;16(6):671-692 [FREE Full text] [Medline: 11780707]
13. Pennbridge J, Moya R, Rodrigues L. Questionnaire survey of California consumers' use and rating of sources of health care information including the Internet. West J Med 1999;171(5-6):302-305 [FREE Full text] [Medline: 10639865]
14. Fox S, Jones S. The social life of health information. Washington, DC: Pew Internet & American Life Project; 2011 May 12. URL: http://www.pewinternet.org/2011/05/12/the-social-life-of-health-information-2011/ [accessed 2014-05-09] [WebCite Cache ID 6PRUon8Xi]
15. California HealthCare Foundation. The State of Health Information Technology in California. Oakland, CA: California HealthCare Foundation; 2008.
16. Couper MP, Singer E, Levin CA, Fowler FJ, Fagerlin A, Zikmund-Fisher BJ. Use of the Internet and ratings of information sources for medical decisions: results from the DECISIONS survey. Medical Decision Making 2010;30:106-114. [doi: 10.1177/0272989X10377661]
17. Gombault V. Deux ménages sur trois disposent d'internet chez eux. Insee Premiere n° 1340 2011:1-4.
18. Renahy E. Recherche d'information en matière de santé sur internet: déterminants, pratiques et impact sur la santé et le recours au soin. Paris: Thèse de doctorat de l'Université Pierre et Marie Curie 2008; 2008:274.
19. Renahy E, Parizot I, Chauvin P. Determinants of the frequency of online health information seeking: results of a web-based survey conducted in France in 2007. Inform Health Soc Care 2010 Jan;35(1):25-39 [FREE Full text] [doi: 10.3109/17538150903358784] [Medline: 20302437]
20. Beck F, Gautier A, Guignard R, Richard JB. Méthode d'enquête du Baromètre santé. In: Beck F, Richard JB, editors. Les comportements de santé des jeunes Analyses du Baromètre santé. Saint-Denis: Inpes; 2010:27-55.
21. Kish L. A procedure for objective respondent selection within the household. JASA 1949;44:380-387.
22. Beck F, Legleye S, Peretti-Watel P. Using the telephone in general population surveys on drugs. In: Decorte T, Korf D, editors. European Studies on Drugs and Drug Policy. Brussels: VUBPress; 2004:113-140.
23. Beck F, Gautier A, Guignard R, Richard JB. Une méthode de prise en compte du dégroupage total dans le plan de sondage des enquêtes téléphoniques auprès des ménages. In: Tremblay ME, Lavallée P, El Hadj Tirari M, editors. Pratiques et Méthodes de sondage. Paris: Dunod, Collection Sciences Sup; 2011:310-314.
24. Beck F. Health Barometer 2010 Methods. 2014. URL: http://www.inpes.sante.fr/Barometres/barometre-sante-2010/pdf/Health-Barometer-2010-Methods.pdf [accessed 2014-04-17] [WebCite Cache ID 6Ou0bAkgJ]
25. Parkerson GR, Broadhead WE, Tse CK. The Duke Health Profile. A 17-item measure of health and dysfunction. Med Care 1990 Nov;28(11):1056-1072. [Medline: 2250492]
26. Leplège A, Ecosse E, Verdier A, Perneger TV. The French SF-36 Health Survey: translation, cultural adaptation and preliminary psychometric evaluation. J Clin Epidemiol 1998 Nov;51(11):1013-1023. [Medline: 9817119]

27. Andreassen HK, Bujnowska-Fedak MM, Chronaki CE, Dumitru RC, Pudule I, Santana S, et al. European citizens' use of E-health services: a study of seven countries. BMC Public Health 2007;7:53 [FREE Full text] [doi: 10.1186/1471-2458-7-53] [Medline: 17425798]
28. Kummervold PE, Chronaki CE, Lausen B, Prokosch HU, Rasmussen J, Santana S, et al. eHealth trends in Europe 2005-2007: a population-based survey. J Med Internet Res 2008;10(4):e42 [FREE Full text] [doi: 10.2196/jmir.1023] [Medline: 19017584]
29. Siliquini R, Ceruti M, Lovato E, Bert F, Bruno S, De Vito E, et al. Surfing the internet for health information: an italian survey on use and population choices. BMC Med Inform Decis Mak 2011;11:21 [FREE Full text] [doi: 10.1186/1472-6947-11-21] [Medline: 21470435]
30. Fox S, Duggan M. Health Online 2013. Washington, DC: Pew Internet & American Life Project; 2013 Jan 15. URL: http://pewinternet.org/~/media//Files/Reports/PIP_HealthOnline.pdf [accessed 2014-01-29] [WebCite Cache ID 6Mz86qGWy]
31. Takahashi Y, Ohura T, Ishizaki T, Okamoto S, Miki K, Naito M, et al. Internet use for health-related information via personal computers and cell phones in Japan: a cross-sectional population-based survey. J Med Internet Res 2011;13(4):e110 [FREE Full text] [doi: 10.2196/jmir.1796] [Medline: 22169526]
32. Moretti FA, Oliveira VE, Silva EM. Access to health information on the internet: a public health issue? Rev Assoc Med Bras 2012;58(6):650-658 [FREE Full text] [Medline: 23250092]
33. AlGhamdi KM, Moussa NA. Internet use by the public to search for health-related information. Int J Med Inform 2012 Jun;81(6):363-373. [doi: 10.1016/j.ijmedinf.2011.12.004] [Medline: 22217800]
34. Aliaga C. Les femmes plus attentives à leur santé que les hommes. Insee Premiere 2002:1-4.
35. Dickerson SS. Women's use of the Internet: what nurses need to know. J Obstet Gynecol Neonatal Nurs 2006;35(1):151-156. [doi: 10.1111/j.1552-6909.2006.00004.x] [Medline: 16466365]
36. Powell J, Inglis N, Ronnie J, Large S. The characteristics and motivations of online health information seekers: cross-sectional survey and qualitative interview study. J Med Internet Res 2011;13(1):e20 [FREE Full text] [doi: 10.2196/jmir.1600] [Medline: 21345783]
37. Korp P. Health on the Internet: implications for health promotion. Health Educ Res 2006 Feb;21(1):78-86 [FREE Full text] [doi: 10.1093/her/cyh043] [Medline: 15994845]
38. Gilmour JA. Reducing disparities in the access and use of Internet health information. a discussion paper. Int J Nurs Stud 2007 Sep;44(7):1270-1278. [doi: 10.1016/j.ijnurstu.2006.05.007] [Medline: 16828775]
39. Eng TR, Maxfield A, Patrick K, Deering MJ, Ratzan SC, Gustafson DH. Access to health information and support: a public highway or a private road? JAMA 1998 Oct 21;280(15):1371-1375. [Medline: 9794322]
40. Hart JT. The inverse care law. Lancet 1971;297:405-412.
41. Murray E, Lo B, Pollack L, Donelan K, Catania J, Lee K, et al. The impact of health information on the Internet on health care and the physician-patient relationship: national U.S. survey among 1.050 U.S. physicians. J Med Internet Res 2003;5(3):e17 [FREE Full text] [doi: 10.2196/jmir.5.3.e17] [Medline: 14517108]
42. Ayers SL, Kronenfeld JJ. Chronic illness and health-seeking information on the Internet. Health (London) 2007 Jul;11(3):327-347. [doi: 10.1177/1363459307077547] [Medline: 17606698]
43. Burns JR, Rapee RM. Adolescent mental health literacy: young people's knowledge of depression and help seeking. J Adolesc 2006 Apr;29(2):225-239. [doi: 10.1016/j.adolescence.2005.05.004] [Medline: 15996727]
44. Bansil P, Keenan NL, Zlot AI, Gilliland JC. Health-related information on the Web: results from the HealthStyles Survey, 2002-2003. Prev Chronic Dis 2006 Apr;3(2):A36 [FREE Full text] [Medline: 16539777]
45. Bundorf MK, Wagner TH, Singer SJ, Baker LC. Who searches the internet for health information? Health Serv Res 2006 Jun;41(3 Pt 1):819-836 [FREE Full text] [doi: 10.1111/j.1475-6773.2006.00510.x] [Medline: 16704514]
46. Couper MP, Kapteyn A, Schonlau M, Winter J. Noncoverage and nonresponse in an Internet survey. Social Science Research 2007 Mar;36(1):131-148. [doi: 10.1016/j.ssresearch.2005.10.002]
47. Hargittai E. First Monday. 2002. Second-level digital divide URL: http://firstmonday.org/article/view/942/864 [accessed 2014-04-17] [WebCite Cache ID 6Ou1wOdmw]
48. Katz JE, Rice RE. Social Consequences of Internet Use: Access, Involvement, and Interaction. Cambridge, MA: MIT Press; 2002:1-463.
49. Chauvin P, Parizot I. Santé et expériences de soins: de l'individu à l'environnement social. Paris: Editions Inserm-Vuibert; 2005:2005.
50. 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]
51. Pappachan MJ. Increasing prevalence of lifestyle diseases: high time for action. Indian J Med Res 2011 Aug;134:143-145 [FREE Full text] [Medline: 21911964]
52. Houston TK, Allison JJ. Users of Internet health information: differences by health status. J Med Internet Res 2002;4(2):e7. [doi: 10.2196/jmir.4.2.e7]
53. Renahy E, Parizot I, Chauvin P. Health information seeking on the Internet: a double divide? Results from a representative survey in the Paris metropolitan area, France, 2005-2006. BMC Public Health 2008;8:69 [FREE Full text] [doi: 10.1186/1471-2458-8-69] [Medline: 18291024]

54. Bianco A, Zucco R, Nobile CG, Pileggi C, Pavia M. Parents seeking health-related information on the Internet: cross-sectional study. J Med Internet Res 2013;15(9):e204 [FREE Full text] [doi: 10.2196/jmir.2752] [Medline: 24047937]
55. Pandey SK, Hart JJ, Tiwary S. Women's health and the internet: understanding emerging trends and implications. Soc Sci Med 2003 Jan;56(1):179-191. [Medline: 12435560]
56. Chambers D, Murphy F. Learning to Reach Out: Young People, Mental Health Literacy and the Internet. Dublin: Inspire Ireland Foundation; 2011.
57. Rietmeijer CA, Shamos SJ. HIV and sexually transmitted infection prevention online: Current state and future prospects. Sex Res Soc Policy 2007;4:65-73.
58. 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]
59. Bailey JV, Murray E, Rait G, Mercer CH, Morris RW, Peacock R, et al. Interactive computer-based interventions for sexual health promotion. Cochrane Database Syst Rev 2010;9:81. [doi: 10.1002/14651858.CD006483.pub2]
60. Grant BF, Dawson DA. Introduction to the national epidemiologic survey on alcohol and related conditions. Alcohol Health Res World 2006;29:74.
61. INPES. URL: http://www.attraction-lemanga.fr/site/index.php [accessed 2014-04-17] [WebCite Cache ID 6Ou2ZSZO5]

Abbreviations

CATI: computer-assisted telephone interview
CNIL: Commission Nationale de l'Informatique et des Libertés
HAS: Haute Autorité de Santé
HON: Health on the Net
INPES: Institut National de Prévention et d'Education pour la Santé
INSEE: Institut National de la Statistique et des Etudes Economiques
MH5: 5-item Mental Health scale from the SF-36
SF-36: 36-item Short-Form health questionnaire

Edited by G Eysenbach; submitted 24.09.13; peer-reviewed by P Singh, T van Mierlo; comments to author 30.10.13; revised version received 12.12.13; accepted 23.01.14; published 13.05.14.

Please cite as:
Beck F, Richard JB, Nguyen-Thanh V, Montagni I, Parizot I, Renahy E
Use of the Internet as a Health Information Resource Among French Young Adults: Results From a Nationally Representative Survey
J Med Internet Res 2014;16(5):e128
URL: http://www.jmir.org/2014/5/e128/
doi:10.2196/jmir.2934
PMID:24824164

Original Paper

Pain-QuILT: Clinical Feasibility of a Web-Based Visual Pain Assessment Tool in Adults With Chronic Pain

Chitra Lalloo[1], BHSc(Hons); Dinesh Kumbhare[2], MD, MSc; Jennifer N Stinson[3], RN-EC, CPNP, PhD; James L Henry[4], PhD

[1]Medical Sciences Graduate Program, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada

[2]Toronto Rehabilitation Institute, Division of Physical Medicine and Rehabilitation, University of Toronto, Toronto, ON, Canada

[3]Department of Child Health Evaluative Sciences, The Hospital for Sick Children, Lawrence S Bloomberg, Faculty of Nursing, University of Toronto, Toronto, ON, Canada

[4]Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada

Corresponding Author:
Chitra Lalloo, BHSc(Hons)
Medical Sciences Graduate Program
Faculty of Health Sciences
McMaster University
HSC-4N35
1280 Main Street West
Hamilton, ON, L8S 4L8
Canada
Phone: 1 905 525 9140
Fax: 1 905 389 3563
Email: lallooc@mcmaster.ca

Abstract

Background: Chronic pain is a prevalent and debilitating problem. Accurate and timely pain assessment is critical to pain management. In particular, pain needs to be consistently tracked over time in order to gauge the effectiveness of different treatments. In current clinical practice, paper-based questionnaires are the norm for pain assessment. However, these methods are not conducive to capturing or tracking the complex sensations of chronic pain. Pain-QuILT (previously called the Iconic Pain Assessment Tool) is a Web-based tool for the visual self-report and tracking of pain (quality, intensity, location, tracker) in the form of time-stamped records. It has been iteratively developed and evaluated in adolescents and adults with chronic pain, including usability testing and content validation. Clinical feasibility is an important stepping-stone toward widespread implementation of a new tool. Our group has demonstrated Pain-QuILT clinical feasibility in the context of a pediatric chronic pain clinic. We sought to extend these findings by evaluating Pain-QuILT clinical feasibility from the perspective of adults with chronic pain, in comparison with standard paper-based methods (McGill Pain Questionnaire [MPQ] and Brief Pain Inventory [BPI]).

Objective: The goal of our study was to assess Pain-QuILT for (1) ease of use, (2) time for completion, (3) patient preferences, and (4) to explore the patterns of self-reported pain across the Pain-QuILT, MPQ, and BPI.

Methods: Participants were recruited during a scheduled follow-up visit at a hospital-affiliated pain management and physical rehabilitation clinic in southwestern Ontario. Participants self-reported their current pain using the Pain-QuILT, MPQ, and BPI (randomized order). A semistructured interview format was used to capture participant preferences for pain self-report.

Results: The sample consisted of 50 adults (54% female, 27/50) with a mean age of 50 years. Pain-QuILT was rated as significantly easier to use than both the MPQ and BPI (P<.01) and was also associated with the fewest difficulties in completion. On average, the time to complete each tool was less than 5 minutes. A majority of participants (58%, 29/50) preferred Pain-QuILT for reporting their pain over alternate methods (16%, 8/50 for MPQ; 14%, 7/50 for BPI; 12%, 6/50 for “other”). The most commonly chosen pain descriptors on MPQ were matched with Pain-QuILT across 91% of categories. There was a moderate-to-high correlation between Pain-QuILT and BPI scores for pain intensity (r=.70, P<.01).

Conclusions: The results of this clinical feasibility study in adults with chronic pain are consistent with our previously published pediatric findings. Specifically, data indicate that Pain-QuILT is (1) easy to use, (2) quick to complete, (3) preferred by a majority of patients, and (4) correlated as expected with validated pain measures. As a digital, patient-friendly method of assessing and

tracking pain, we conclude that Pain-QuILT has potential to add significant value as one standard component of chronic pain management.

(J Med Internet Res 2014;16(5):e127) doi:10.2196/jmir.3292

KEYWORDS

chronic pain; assessment tool; Internet; clinical feasibility

Introduction

Chronic pain, defined as pain that persists beyond normal time of healing, is a prevalent and debilitating problem that is now recognized as a disease [1-3]. Common types of chronic pain include low back, headache, abdominal, musculoskeletal, and neuropathic pain [4]. Pain is a complex sensory and emotional phenomenon that, while intensely experienced, is often difficult to communicate [5].

Accurate and timely pain assessment is critical to developing and monitoring a pain management plan [6]. Given that there is no medical test to directly measure pain, health care providers rely primarily on patient self-report, including pain quality (what it feels like), intensity (how much it hurts), location (spatial distribution), and temporal nature (how it changes over time) [6]. Assessment of pain quality and location is particularly important because this information can be used to distinguish between different diagnostic subgroups (eg, neuropathic versus non-neuropathic pain) [7,8].

Chronic pain management often takes place across multiple settings (eg, hospitals, clinics) and involves numerous health care providers, including physicians, nurses, physiotherapists, chiropractors, and psychologists [9-11]. Pain outcomes need to be consistently tracked over time in order to gauge the effectiveness of different management strategies, including physical, psychological, and pharmacological approaches. However, there is often a lack of consistency in the assessment of pain across these different settings and providers. One reason for this lack of consistency is the standard use of paper-based assessment tools, which are not conducive to tracking pain over time. Commonly used paper-based tools include the McGill Pain Questionnaire (MPQ) [12] and the Brief Pain Inventory (BPI) [13]. However, there is limited research on the serial use of these measures in clinical pain assessment.

The emergence of Internet and mobile technology has created opportunities for innovation in the field of pain assessment and management. For example, electronic pain diaries offer advantages such as ease of data tracking, improved patient compliance, and capture of real-time pain reports without memory bias [14-20].

Pain-QuILT is a Web-based tool for the visual self-report and tracking of pain (quality, intensity, location, tracker) in the form of time-stamped records [21-24]. Pain quality is expressed by choosing from a validated library of labeled pain icons, such as a matchstick for “burning pain”. Pain intensity is quantified using a 0-10 numerical rating scale (NRS) ranging from “no pain” to “worst pain imaginable”. Pain location is illustrated by “dragging-and-dropping” pain icons onto a detailed virtual body-map that is codified into over 100 regions.

To our knowledge, Pain-QuILT is the only tool that captures the complex sensations of chronic pain by allowing patients to self-report different qualities and intensities of pain across their entire body. For example, they can record the simultaneous experience of a “3/10” burning pain in their shoulder as well as a “5/10” pain in their foot that is both “burning” and “sharp”. All reported data are digitally captured and then populated into a database, which can be used to track changes in pain quality and intensity across different body regions over time.

Health care professionals can use this information to monitor the effectiveness of any pain management practices. Patients can keep track of their pain to help inform self-management in the home setting. By standardizing the assessment of pain outcomes in a digitized format, Pain-QuILT may also improve the coordination of pain management across multiple health care providers.

Pain-QuILT has been iteratively developed and evaluated in adolescents and adults with chronic pain, including usability testing and content validation. Before widespread implementation of Pain-QuILT, it is critical to evaluate clinical feasibility (ie, the ease with which it can be applied in a real-world setting), compared with standard methods of pain assessment. Recently, our group established clinical feasibility in an interdisciplinary pediatric chronic pain clinic that used a semistructured interview method to assess pain [24]. In comparison with this standard method, Pain-QuILT was preferred by a majority of adolescent patients and was perceived to be clinically useful for visually capturing pain and promoting better communication between patients and health providers.

Given that the MPQ and BPI are the standard tools used in adults, the purpose of this study was to extend the findings from our pediatric work to evaluate clinical feasibility of Pain-QuILT among adults with chronic pain in comparison with the MPQ and BPI. In the context of this clinical feasibility study, our primary aims were to assess Pain-QuILT for (1) ease of use, (2) time for completion, and (3) patient preferences. Our secondary aim was to explore the patterns of self-reported pain across the comparator methods of Pain-QuILT, MPQ, and BPI.

Methods

Study Setting

This study was conducted at a hospital-affiliated pain management and physical medicine and rehabilitation outpatient clinic in southwestern Ontario. It was staffed by an interdisciplinary team of health care professionals, consisting of a physiatrist, physical therapist, and kinesiologist. Patients who are referred to this outpatient clinic receive a thorough medical evaluation, including assessment of pain, and are then informed of the management plan including pharmacological,

injection, and physical therapies. They may also be referred for psychological therapy (eg, group counseling, cognitive behavioral therapy) if needed. All patients are reassessed at timely intervals and treatments are adjusted according to clinical need.

Recruitment

Informed written consent was obtained from all participants, and the study was approved by the locally responsible Research Ethics Boards. A health care provider known to patients identified eligible individuals by screening the patient lists of consecutively scheduled clinic appointments. Individuals were eligible to participate if they were (1) aged 18 years or older, (2) able to speak and read English, and (3) currently experiencing pain of any intensity according to self-report. Individuals were excluded if they had severe cognitive impairment or major comorbid medical or psychiatric illness that could preclude their ability to self-report pain or take part in a verbal interview, according to their health care provider. Individuals were also excluded if they had severe vision or hand dexterity impairments that could prevent independent use of a computer and mouse.

Demographic and Health-Related Data

Following consent, each participant completed a Demographic and Health Questionnaire, which collected data on age, sex, computer comfort, weekly computer use, language proficiencies, education level, and date of pain problem onset.

Interview Protocol

All participants took part in an individual semistructured interview (20-30 minutes) with a trained investigator (author CL). The investigator was experienced in conducting qualitative interviews and used techniques to minimize the power differential between the interviewer and participant (eg, established rapport, engaged in active listening, used relaxed body language) [25]. The investigator also stressed that the research team wished to ensure that Pain-QuILT addressed the needs of adults with chronic pain and thus encouraged participants to freely express opinions about good and bad aspects of the tool. As a first step, participants self-reported their pain using the Pain-QuILT, MPQ, and BPI (described in detail below). These tools were administered in a randomized order for each participant, in order to minimize potential order effects. Investigator observation and participant comments were used to identify any difficulties or confusion with using each tool; these were recorded as field notes. The time required to complete each tool was recorded. Next, a semistructured interview format was used to discuss participant preferences for pain self-report. A 0-10 NRS ranging from “not easy at all” to “very easy” was used to appraise each tool. Qualitative written feedback on the ease of using each tool was also collected. Finally, participants were asked to indicate their preference of methods for self-reporting pain and explain the reason for their choice. All interviews were conducted by the same investigator in a quiet room within the clinic.

Pain Tool Comparison

McGill Pain Questionnaire

This paper-based questionnaire was developed in the 1970s through groundbreaking research that was focused on identifying common word descriptors for the pain experience [12,26,27]. At the time, there was no available tool that accounted for the multidimensional nature of pain. The MPQ is composed of 20 subclasses that correspond to sensory, affective, evaluative, and miscellaneous pain. Each subclass consists of a clustered list of 2-6 word descriptors. For example, the first subclass of word descriptors is “sensory-temporal” and is made up of the descriptors: “flickering; quivering; pulsing; throbbing; beating; pounding”. There is a total of 78 descriptors on the MPQ. Participants were instructed to review each discrete cluster of words and then select the one word that best described their current pain. If none of the words within a cluster were descriptive of their pain, then no word was selected. The MPQ is one page in length and was administered by the study investigator.

Brief Pain Inventory Short Form

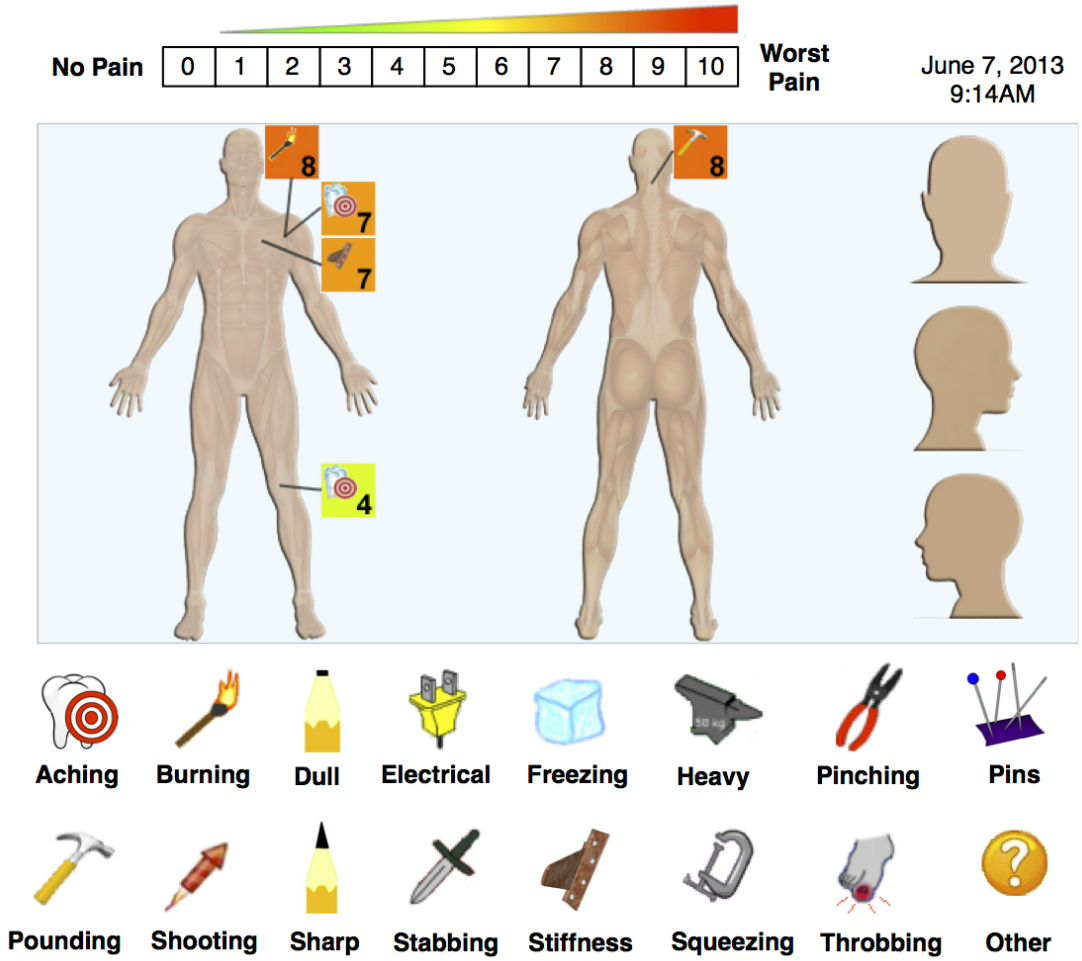
This paper-based questionnaire was developed in the 1980s for patients with cancer pain, based on research suggesting that existing measures such as the MPQ were burdensome for patients to complete [13,28]. Since its initial development, the BPI has subsequently become one of the most widely used tools for assessing all types of pain in both clinical and research settings [29]. It is designed to assess pain location and severity as well as level of interference with daily life. In the present study, participants used a pen to shade painful areas on a body-manikin diagram. The body-manikin consisted of anterior and posterior aspects and included no regional demarcations. Next, participants were required to rate the intensity of their “pain right now” as well as their “worst”, “least”, and “average” pain from the past 24 hours using separate 0-10 NRS items ranging from “no pain” to “pain as bad as you can imagine”. Finally, participants were asked to rate the extent to which pain had interfered with different parts of their life in the past 24 hours. Each quality of life domain was rated on a separate NRS ranging from 0 (“does not interfere”) to 10 (“completely interferes”). The BPI is one page in length and was administered by the study investigator.

Pain-QuILT

Participants were taught how to use Pain-QuILT via a standard 3-minute demonstration. Following confirmation of understanding, each participant was instructed to use the investigator laptop computer (MacBook Pro) with external mouse to “create a picture” of their current pain, as illustrated in Figure 1. First, they chose from the library of labeled pain quality icons to describe what their pain felt like. The Pain-QuILT library consisted of 16 icons to represent aching, burning, dull, electrical, freezing, heavy, pinching, pins & needles, pounding, shooting, sharp, stabbing, stiffness, squeezing, throbbing, and “other” pain. They then used the mouse to “drag-and-drop” a miniature copy of this descriptive icon onto a virtual body-map to show pain location. The entire body-map was displayed on a single screen and was made up

of anterior and posterior aspects, as well as magnified views of the head (anterior, posterior, side-view). The body-map was codified into 110 distinct regions, and each region became highlighted in blue as the computer mouse hovered over it. Next, after "dropping" the icon onto the appropriate body region, the user assigned a rating of intensity for this pain by using a "pop-up" 0-10 NRS ranging from "no pain" to "worst pain imaginable". The 0-10 NRS also corresponded with a color scale ranging from green (lower intensity) to red (higher intensity). After the user had chosen an intensity value, the pain icon was added to the body-map, along with the numerical rating. The dropped icon-number pair was enclosed within a square box whose fill color corresponded to the intensity rating (eg, dark green fill for a rating of 1/10). Users continued to "drag-and-drop" numbered icons onto the virtual body-map until all of their current pain or pains had been recorded. Figure 1 shows a patient reporting multiple pains across their body of different qualities and intensities, specifically, shoulder pain that is both "burning" and "aching", a painful stiffness in their chest, an "aching" knee pain, and a "pounding" pain in the back of their neck. All user-entered pain data (quality, intensity, location), as well as information on time and date of entry, were automatically uploaded to a back-end database that was accessible to the research team.

Figure 1. Screenshot of Pain-QuILT user interface for self-reporting the quality, intensity, and location of current pain. Copyright McMaster University. Used with permission. All permission requests for this image should be made to the copyright holder (McMaster Industry Liaison Office).



Data Analysis

Qualitative written data and field notes from the semistructured interview were transcribed verbatim and imported into the qualitative software program, HyperRESEARCH [30]. This software was used to facilitate a simple content analysis of the data [31]. A line-by-line coding analysis was used to identify key concepts from the interview transcripts and field notes. Concepts addressed during the semistructured interviews were used to thematically code and organize participant responses [31]. Participant quotations were selected to illustrate each key interview concept with the aim of representing the balance of opinion among participants.

Quantitative data from the Demographic and Health Questionnaire, MPQ, BPI, and Pain-QuILT were coded, scored, and entered into a Statistical Package for the Social Sciences database [32]. As described by Lalloo and colleagues, the extracted parameters from each Pain-QuILT report were the number of unique painful sites (range 0 to 110) and number of different pain quality descriptors (range 0 to 16) used to express current pain [24]. Additionally, a cumulative mean pain intensity score was calculated across all painful body sites. While this

cumulative score provided a convenient indicator of the central tendency of data, it was also sensitive to outliers. Thus, we also extracted the lowest and highest single NRS intensity score to provide an indicator of data dispersion. For example, if a participant reported a 5/10 burning pain in their foot, a 3/10 burning pain in their hand, and a 3/10 stabbing pain in their back, then the number of unique painful sites would be recorded as 3, the number of unique pain quality descriptors would be 2, the cumulative intensity score would be calculated as [(5+3+3)/3]=3.7, the lowest reported NRS score would be 3, and the highest reported NRS score would be 5.

All data were analyzed descriptively to assess measures of central tendency (mean, median) and dispersion [standard deviation, interquartile range]. Data were also evaluated to ensure that they met the assumptions of parametric statistical analysis (ie, the normal distribution). When these assumptions were not met, the non-parametric equivalent test was used. Repeated measure analysis of variance (ANOVA) was used to determine whether there were any differences between Pain-QuILT, MPQ, and BPI in terms of time to complete or ease of use ratings. Pearson correlations were used to examine the association between pain intensity scores on Pain-QuILT and BPI. The *a priori* criterion for evidence of convergent validity was a moderate correlation of r=.5 between Pain-QuILT and BPI scores for current pain intensity. Using the guidelines from Streiner and Norman pertaining to sample size for correlation coefficients, assuming alpha=.05 and beta=.05, the required sample size was N=50 [33]. The level of significance was set at P<.05 for all tests.

Results

Participant Characteristics

A total of 50 adults completed the study over a 5-month period in 2013. Sample characteristics are summarized in Table 1. Nearly all participants (48/50, 96%) had a computer at home as well as Internet access (45/50, 90%). Of the 50 participants, 84% (42/50) reported being “comfortable” or “very comfortable” with using computers, while 10% (5/50) were “a little comfortable” and 6% (3/50) were “not at all comfortable”. The self-reported frequency of computer use among participants was none (3/50, 6%), once per week (3/50, 6%), twice per week (2/50, 4%), three times per week (4/50, 8%), five times per week (1/50, 2%), and every day (37/50, 74%).

Table 1. Characteristics of study participants (N=50).

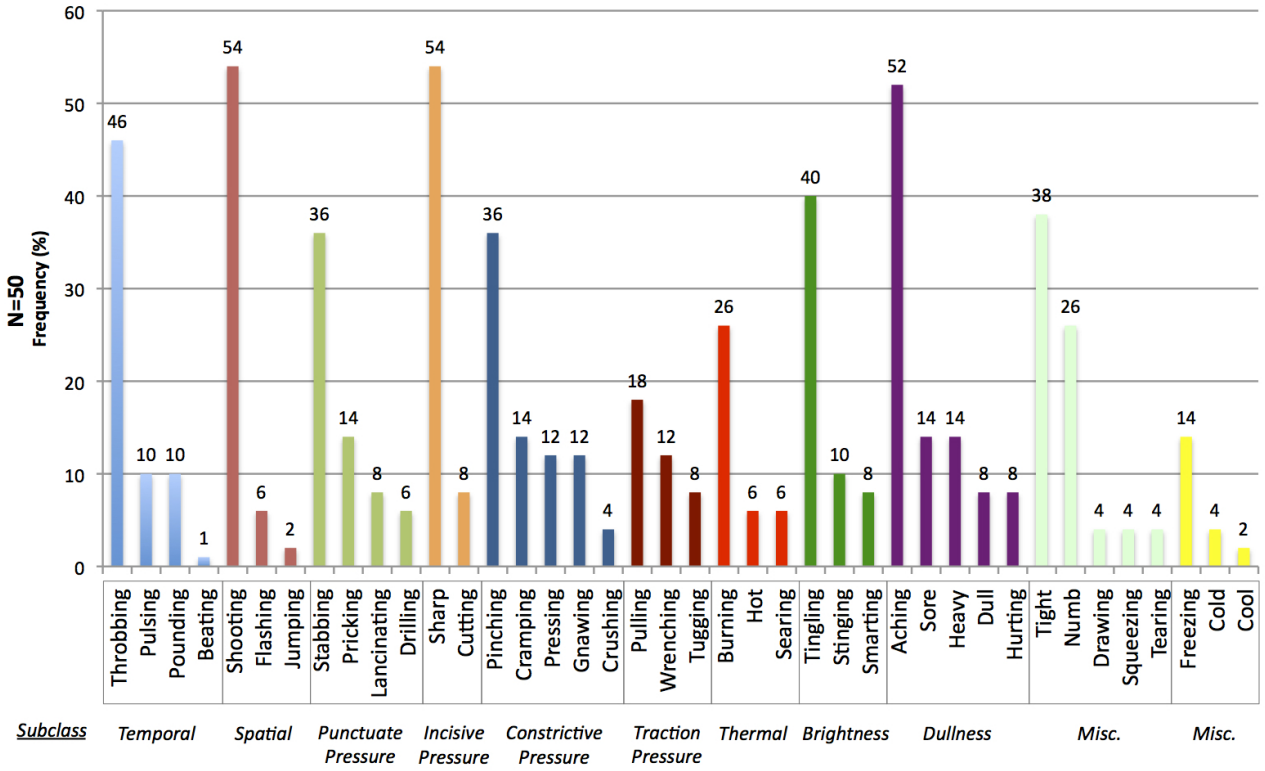
Characteristics	n
Age in years, mean (SD) (range)	50 (14) (18-76)
Gender, n (%)	
Male	23 (46)
Female	27 (54)
Language, n (%)	
English as first spoken language	39 (78)
Spoke English only	31 (62)
Spoke English and another language	19 (38)
Total years education, mean (SD) (range)	13.8 (3.8) (0-21)
Chronic pain duration in years, mean (SD) (range)	8.3 (8.9) (1-33)
Current pain treatment modalities, n (%)	
Pharmacological	43 (86)
Physical therapy	19 (38)
Massage therapy	9 (18)
Alternative or complementary	5 (10)
Chiropractic therapy	2 (4)
Acupuncture	2 (4)
Pain interference in past 24 hours, mean (SD)	
Normal work	7.2 (2.5)
Enjoyment of life	7.0 (2.9)
Sleep	6.8 (2.6)
General activity	6.7 (2.5)
Mood	6.2 (2.9)
Walking ability	5.6 (3.1)
Relations with other people	5.0 (2.9)

Self-Reported Pain

McGill Pain Questionnaire

The relative endorsement of MPQ pain quality descriptors between and within subclasses is illustrated in Figure 2. The most commonly chosen MPQ words to express current pain were matched with a descriptor in the Pain-QuILT library across all subclasses, except for "miscellaneous". This pattern was consistent regardless of whether the MPQ was administered before Pain-QuILT (29/50, 58%), or Pain-QuILT was administered before the MPQ (21/50, 42%).

Figure 2. Relative frequency of words chosen by participants on the McGill Pain Questionnaire to describe their current pain.



Brief Pain Inventory

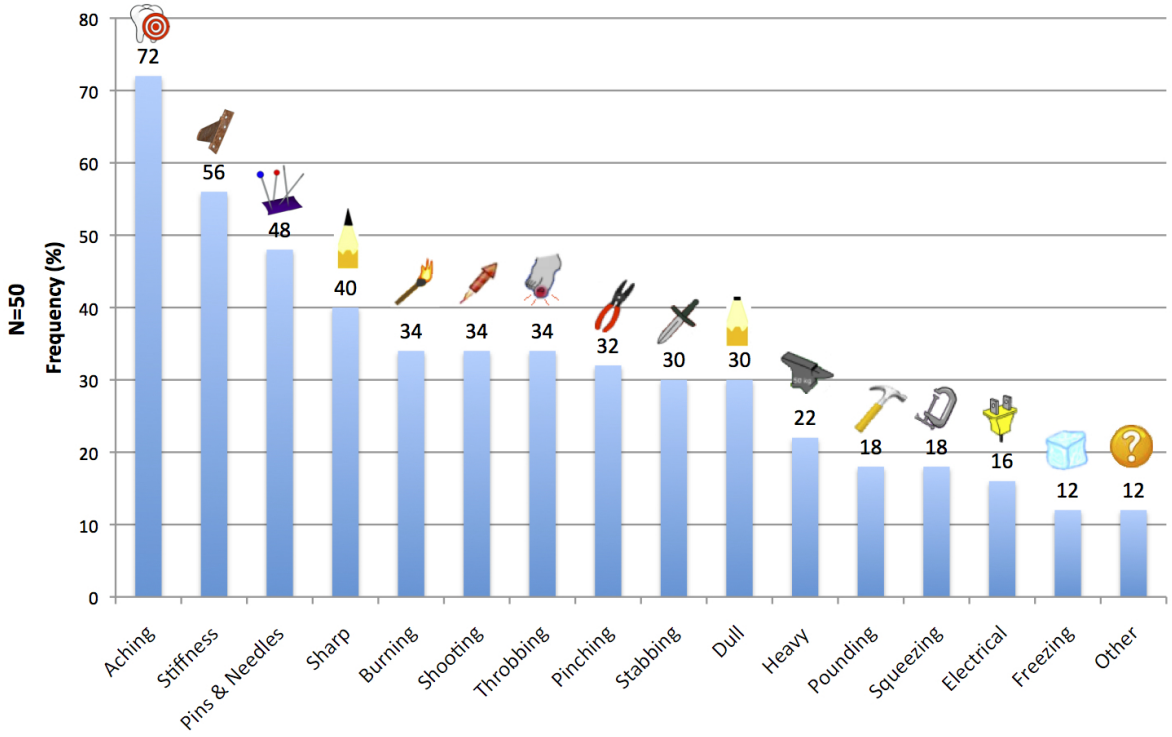
The mean score for current reported pain intensity was 6.6 (SD 2.1). The mean scores for recalled pain in the past 24 hours were 7.9 (SD 1.4) for "worst" pain, and 4.4 (SD 2.2) for "least" pain, respectively.

Pain-QuILT

The mean number of unique painful sites reported was 6.5 (SD 4.0, range 1-22). The mean number of different pain qualities used to describe current pain was 5.0 (SD 2.4, range 1-10). The relative endorsement of *Pain-QuILT* icons across all participants is illustrated in Figure 3. The mean reported intensity for current pain (ie, the cumulative calculated score across all body sites) was 6.2 (SD 2.0). The mean lowest reported pain intensity score was 4.8 (SD 2.1), and the mean highest reported pain intensity score was 7.4 (SD 2.1).

The Pearson correlation coefficient between the *Pain-QuILT* score for current pain (calculated across all body sites) and BPI score for current pain (single NRS rating) was r=.70. The Pearson correlation coefficient between the highest reported intensity score on Pain-QuILT and the BPI score for current pain was r=.76. The Pearson correlation coefficient between the lowest reported intensity score on Pain-QuILT and the BPI score for current pain was r=.55.

Figure 3. Relative frequency of Pain-QuILT icons chosen by participants to describe their current pain.



Ease of Use

All participants reported the relative ease of using each tool for self-reporting pain. The mean ratings were 5.9 (SD 2.6) for the MPQ, 7.0 (SD 2.6) for the BPI, and 8.3 (SD 2.0) for Pain-QuILT. Overall, there was a significant difference between the tools in terms of perceived ease of use, $F_{2,96}$=20.6, P<.001. Pairwise comparisons also indicated significant differences between the MPQ and BPI (P=.009), MPQ and Pain-QuILT (P<.001), as well as BPI and Pain-QuILT (P=.002).

Participant-Reported Difficulties With Using Each Pain Tool

Overall, 46% (23/50) of participants indicated that they had difficulties in completing the MPQ, while 22% (11/50) reported difficulties with the BPI and 16% (8/50) specified difficulties with Pain-QuILT.

The most commonly reported issue with the MPQ was trouble with understanding the qualitative word descriptors (10/23, 43%) due to language barriers (eg, English as second language) or uncommon vocabulary, such as “taut” and “smarting”. Participants (7/23, 30%) also reported that the available pain words “...weren’t very good to describe [pain]” (ie, lack of descriptiveness). Other participants (7/23, 30%) noted that it was difficult to select the right words to express their pain due to ambiguity (“what is difference between cool, cold, freezing?”), the number of available options (“too many choices”), and the presence of more than one relevant word from certain subclasses. Last, participants (2/23, 9%) expressed concern about potentially misrepresenting their pain to their health care providers: “more fear of not describing your pain properly with this test”.

The most commonly reported issues with the BPI were communicating pain location using the body-manikin (2/11, 18%; “hard to pull out meaning”) and choosing intensity ratings to describe pain (2/11, 18%; “hard time with pain numbers”). Other reported difficulties included recalling pain over the last 24 hours (“hard to simplify pain”), reporting pain from multiple sites (“varying intensities of pain from different injuries”), and questionnaire design (“cumbersome to complete, too general”). One participant also indicated a “fear of not explaining properly what is happening”.

The most commonly reported issues with Pain-QuILT were related to the virtual body-map (3/8, 37.5%). Specifically, participants identified a need for orientation labels (left, right) and to make it easier to isolate specific painful body areas (“hard to find specific regions on [the] back versus a ‘paint’ tool, because some pain radiates”). In addition, participants (2/8, 25%) indicated difficulty in choosing pain quality icons due to “too many choices...sometimes it aches, sometimes it burns”, and a dislike of using descriptors because “pain just hurts”. Other participants (3/8, 37.5%) identified a “bug” in the software related to an inability to remove icons that were mistakenly added to the body-map.

Time to Complete

The mean time required by participants to complete a single pain report using each tool was 4.2 minutes (SD 1.5) for the MPQ, 4.0 minutes (SD 1.4) for the BPI, and 4.1 minutes (SD 2.2) for Pain-QuILT. There was no significant difference between the tools in terms of time to complete, $F_{1.4,44.8}$=0.13, P=.81.

Participant Preferences for Self-Reporting Pain

Overall, 16% (8/50) participants chose the MPQ as their preferred method for self-reporting pain, while 14% (7/50) chose the BPI, and 58% (29/50) chose Pain-QuILT. Four of the 50 participants (8%) indicated that they preferred the "other" method of verbally explaining pain to their health care provider. Finally, one participant (1/50, 2%) indicated an equal preference between the BPI and Pain-QuILT.

Reasons for selecting the MPQ included a preference for paper versus electronic pain reporting and greater perceived precision in describing pain, for example, "[it has] words that exactly indicate what is happening to [my] leg—bang on".

Reasons for choosing the BPI included familiarity, the ability to describe how pain changes over time, and ease of choosing ratings on a set scale, for example, "[it] seems more easy to answer personally. Fits the way that I speak".

Explanations for choosing Pain-QuILT included greater ease of use, ability to pinpoint different locations and types of pain, preference for computer versus paper-based pain reporting, as well as the visual language to express pain, for example, "[I would] feel more confident being treated by a doctor if they used this tool because [they] would know exactly what you are feeling".

Discussion

Previous Findings

Our previous work has established the acceptability, usability, and content validity of Pain-QuILT in samples of adults with central post-stroke pain [21], adults in the community with a range of different types of chronic pain [22], as well as adults and adolescents with arthritis pain [34]. Clinical feasibility testing, the focus of the present study, is an important stepping-stone toward widespread implementation of a new assessment tool [35]. Our group has recently demonstrated clinical feasibility of Pain-QuILT in the context of an interdisciplinary pediatric chronic pain clinic among adolescents aged 12-18 years [24]. The present study sought to extend these findings by evaluating clinical feasibility of Pain-QuILT from the perspective of adults attending an outpatient pain clinic for treatment of chronic pain. This study included a comparison of Pain-QuILT with standard methods of pain assessment.

Principal Results

As a tool for self-reporting pain, Pain-QuILT was rated as significantly easier to use than the MPQ and BPI, which are two of the most commonly used pain assessment tools in research and clinical practice. Almost half (46%) of participants reported difficulties in using the MPQ. Most of these difficulties related to understanding the pain descriptors and finding accurate words to express pain from a large number of options. These findings of the present study are consistent with a meta-analysis of 51 studies involving 3624 patients, which found that most MPQ words (75%) are rarely endorsed by patients to describe their pain [36]. Although the BPI was associated with fewer reported difficulties, participants indicated that its design was not conducive to reporting different intensities of pain in different body sites. Numerous studies have demonstrated that chronic pain is rarely confined to a single body region [37-39]. For instance, in a study involving 2445 patients, Carnes and colleagues found that 73% experienced pain across multiple body sites [37]. Among patients with low back pain, only 13% experienced regionally isolated pain. In terms of implications for pain assessment and management, these authors concluded, "self-reported measures of multi-site pain are problematic with pain measures that are site-specific. Pain in other areas may render them less reliable and responsive. Future intervention studies should consider recording other pain sites to identify predictors of response to treatment" (p. 1170) [37]. Overall, Pain-QuILT was associated with the fewest reported difficulties among participants. Most of the identified issues (75%) will be resolved in the next iteration of Pain-QuILT software (eg, adding orientation labels to body map, fixing "bug" related to deleting unwanted icons). Participant concerns related to the changing nature of pain ("sometimes it aches, sometimes it burns") will be addressed in future longitudinal studies, which will allow patients to use Pain-QuILT as a diary to document symptoms as they occur. A major identified strength of Pain-QuILT was the ability to record multiple sites, types, and intensities of pain.

The average time required to complete each assessment tool was less than 5 minutes. While there was no significant time difference between the tools, it is important to note that patients can enter Pain-QuILT data independently, while the MPQ and BPI are usually administered by a health care provider in the context of a clinic appointment. Moreover, Pain-QuILT data are generated and stored in a digital format, while information from MPQ and BPI must be manually transcribed into a spreadsheet (paper or computer-based) in order to facilitate tracking over time. Thus, Pain-QuILT has the potential to increase efficiency of clinic appointments by (1) empowering patients to self-report pain on their own time (eg, at home and/or in the clinic waiting room), (2) providing health care providers with digital summaries of tracked pain data to evaluate and inform their management plan, and (3) standardizing the assessment of pain outcomes for use across multiple providers.

Given the inherently personal nature of pain, it is important to consider patient preferences regarding the most effective way of expressing symptoms. The majority of participants (58%) indicated positive preference for Pain-QuILT over alternate methods. It is well recognized that patient engagement is a critical factor in the successful management of chronic disease [40]. In particular, effective doctor-patient communication is known to enhance the health outcomes of pain management [41]. The interactive and dynamic format of Pain-QuILT may also help patients forge a stronger emotional connection to the tool as a means for portraying and conveying their pain experience, compared to static questionnaires. Moreover, there is a growing body of literature documenting the rise of "self-tracking" among people living with chronic illness. A recent Pew Research Center report found that 40% of adults with 1 chronic condition and 62% of adults with 2 chronic conditions currently self-track their symptoms [42]. In terms of patient benefits, respondents indicated that self-tracking influenced their overall approach to maintaining health (56%),

prompted them to ask their doctor new questions (53%), or influenced a treatment decision (45%) [42]. Thus, by providing a user-friendly method for communicating with health care providers and self-tracking painful symptoms, Pain-QuILT may encourage greater patient involvement in the long-term management of their own disease.

There is a growing number of patient-oriented mobile applications (apps) designed to aid the self-tracking of pain. In 2011, Rosser and Eccleston identified 111 pain management apps, of which 24% included a self-monitoring function [43]. A more recent scoping review, conducted in 2013, identified 224 pain apps, of which 14% allowed users to self-track their symptoms [44]. Unfortunately, both studies identified major limitations in the current field of pain apps, including a lack of formal evaluation and limited involvement of health care professionals and patients in their development. Pain-QuILT has been iteratively evaluated and refined through consultation with patients as well as health care professionals and thus has potential to address these identified gaps in the field, as one component of chronic pain management.

Given that participants were asked to self-report their current pain using three different methods, we expected to observe consistency in reported pain. Using the MPQ, participants were presented with a choice of 46 qualitative descriptors across 11 subclasses. Interestingly, the most frequently chosen MPQ words were consistent with the icon descriptors on Pain-QuILT. This relationship was independent of the order of tool assessment. Pain-QuILT icons and word descriptors have been iteratively refined based on patient interviews to ensure that they are representative of the pain experience. The observation that the icons correspond with the most frequently endorsed MPQ descriptors provides further evidence of validity. In terms of pain intensity scores, we examined correlations between Pain-QuILT (body site-specific pain scores) and the BPI (single global score for current pain). There were high correlations ($r \geq .70$) observed between BPI score and (1) the calculated average pain score across all body sites, and (2) the single highest reported pain score across all body sites. There was also a moderate correlation ($r=.55$) observed between BPI score and the single lowest reported pain score across all body sites. Along with our previous pediatric study, which compared Pain-QuILT scores with a verbal NRS ($r=.61$), the current data provides further evidence of convergent validity. Importantly, in terms of clinical usefulness, we suggest that the greater level of detail elicited by Pain-QuILT may help inform pain management strategies (eg, observing how treatment affects pain quality and intensity scores within specific body sites) more than a single global intensity score.

Limitations and Future Directions

The present clinical feasibility study was conducted at a single interdisciplinary pain management and rehabilitation clinic in Southwestern Ontario. Although the organization and treatment model of this site was consistent with other Canadian multidisciplinary pain treatment facilities [45], we acknowledge that future work is needed to evaluate clinical feasibility of Pain-QuILT in other settings. Further, given the interview component of this study, it was necessary for all participants to be able to speak and read English. Although 38% of participants spoke multiple languages, future work is needed to formally evaluate Pain-QuILT in non-English speaking groups. Given the visual nature of Pain-QuILT reporting, it could prove to enhance pain communication for individuals with limited verbal or cognitive skills.

Participants in this study completed only a single Pain-QuILT report. Future work will evaluate whether patient perceptions regarding ease of use and preferences, as well as time to complete, are affected by repeated usage.

Conclusions

The results of this clinical feasibility study in adults with chronic pain are consistent with our previously published pediatric findings [24]. Specifically, data indicate that Pain-QuILT is (1) easy to use, (2) quick to complete, (3) preferred by a majority of adults with chronic pain, and (4) correlated as expected with validated pain measures. As a digital, patient-friendly method of assessing and tracking pain, we conclude that Pain-QuILT has potential to add significant value as one standard component of chronic pain management.

The tool will be licensed for clinical use and research studies through the McMaster Industry Liaison Office [46,47]. Updated information on availability will be provided on the author website [47] and Twitter account (@PainQuILT).

Acknowledgments

PhD work of author CL has received salary support from an Ontario Graduate Scholarship (Ontario Ministry of Training), a PhD Graduate Training Award from The Arthritis Society / Canadian Arthritis Network, and a PhD Research Stipend from the Pain in Child Health Strategic Training Initiative in Health Research (Canadian Institutes of Health Research). Software development of Pain-QuILT has been supported by a Discovery Advancement Project Grant from the Canadian Arthritis Network.

The authors thank Dr David Streiner, PhD, CPsych (Department of Psychiatry & Behavioural Neuroscience, McMaster University) for providing guidance on study design. The authors thank Émilie McMahon-Lacharité, MSc (Biomedical Communications, University of Toronto), creator of the original Iconic Pain Assessment Tool, for supporting this study.

Author CL thanks the members of her graduate supervisory committee at McMaster University: Dr Joy MacDermid, PT, PhD (Department of Rehabilitation Science, McMaster University), Dr Del Harnish, PhD (Department of Pathology, Molecular Medicine & Biology), Dr Kristina Trim, PhD (Bachelor of Health Sciences Program, McMaster University), Dr Jennifer Stinson,

RN, PhD (The Hospital for Sick Children, University of Toronto), and Dr James L Henry, PhD (Department of Psychiatry & Behavioural Neuroscience, McMaster University).

The authors thank Kelsie Korneyi, Navraj Rhandhawa, MD, Jessie Grewal, MSc(kin), and Amy Saward, MSc(kin), for facilitating participant recruitment.

Finally, the authors thank our study participants for their contribution to this research.

Authors' Contributions

All authors contributed substantially to (1) study conception, design, and data interpretation, (2) drafting the manuscript or revising it critically for important intellectual content, and (3) final approval of the version to be published. Participant interviews were conducted by author CL.

Conflicts of Interest

None declared.

References

1. Merskey H, Bogduk N. Classification of chronic pain: Description of chronic pain syndromes and definitions of pain terms. Seattle: IASP Press; 2002.
2. Schopflocher D, Taenzer P, Jovey R. The prevalence of chronic pain in Canada. Pain Res Manag 2011;16(6):445-450 [FREE Full text] [Medline: 22184555]
3. Tracey I, Bushnell MC. How neuroimaging studies have challenged us to rethink: is chronic pain a disease? J Pain 2009 Nov;10(11):1113-1120. [doi: 10.1016/j.jpain.2009.09.001] [Medline: 19878862]
4. Boulanger A, Clark AJ, Squire P, Cui E, Horbay GL. Chronic pain in Canada: have we improved our management of chronic noncancer pain? Pain Res Manag 2007;12(1):39-47 [FREE Full text] [Medline: 17372633]
5. Merskey H, Bogduk N. Part III: Pain Terms, A Current List with Definitions and Notes on Usage. In: Classification of chronic pain: descriptions of chronic pain syndromes and definitions of pain terms. Seattle: IASP Press; 1994.
6. Jensen MP, Karoly P. Self-Report Scales And Procedures For Assessing Pain In Adults. In: Handbook Of Pain Assessment. New York: The Guilford Press; 2001.
7. Jensen MP. The validity and reliability of pain measures in adults with cancer. J Pain 2003 Feb;4(1):2-21. [Medline: 14622723]
8. Hochman JR, French MR, Bermingham SL, Hawker GA. The nerve of osteoarthritis pain. Arthritis Care Res (Hoboken) 2010 Jul;62(7):1019-1023. [doi: 10.1002/acr.20142] [Medline: 20589688]
9. Hasselström J, Liu-Palmgren J, Rasjö-Wrååk G. Prevalence of pain in general practice. Eur J Pain 2002;6(5):375-385. [Medline: 12160512]
10. Todd KH, Ducharme J, Choiniere M, Crandall CS, Fosnocht DE, Homel P, PEMI Study Group. Pain in the emergency department: results of the pain and emergency medicine initiative (PEMI) multicenter study. J Pain 2007 Jun;8(6):460-466. [doi: 10.1016/j.jpain.2006.12.005] [Medline: 17306626]
11. Otis JD, Macdonald A, Dobscha SK. Integration and coordination of pain management in primary care. J Clin Psychol 2006 Nov;62(11):1333-1343. [doi: 10.1002/jclp.20313] [Medline: 16937344]
12. Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. Pain 1975 Sep;1(3):277-299. [Medline: 1235985]
13. Cleeland CS, Ryan KM. Pain assessment: global use of the Brief Pain Inventory. Ann Acad Med Singapore 1994 Mar;23(2):129-138. [Medline: 8080219]
14. Stinson JN, Stevens BJ, Feldman BM, Streiner D, McGrath PJ, Dupuis A, et al. Construct validity of a multidimensional electronic pain diary for adolescents with arthritis. Pain 2008 Jun;136(3):281-292. [doi: 10.1016/j.pain.2007.07.002] [Medline: 17723279]
15. Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient compliance with paper and electronic diaries. Control Clin Trials 2003 Apr;24(2):182-199. [Medline: 12689739]
16. Palmero TM, Valenzuela D. Use of pain diaries to assess recurrent and chronic pain in children. Suffer Child 2003;3:1-19.
17. Jamison RN, Raymond SA, Levine JG, Slawsby EA, Nedeljkovic SS, Katz NP. Electronic diaries for monitoring chronic pain: 1-year validation study. Pain 2001 Apr;91(3):277-285. [Medline: 11275385]
18. Stinson JN, Jibb LA, Nguyen C, Nathan PC, Maloney AM, Dupuis LL, et al. Development and testing of a multidimensional iPhone pain assessment application for adolescents with cancer. J Med Internet Res 2013;15(3):e51 [FREE Full text] [doi: 10.2196/jmir.2350] [Medline: 23475457]
19. Spyridonis F, Ghinea G, Frank AO. Attitudes of patients toward adoption of 3D technology in pain assessment: qualitative perspective. J Med Internet Res 2013;15(4):e55 [FREE Full text] [doi: 10.2196/jmir.2427] [Medline: 23575479]
20. Baggott C, Gibson F, Coll B, Kletter R, Zeltzer P, Miaskowski C. Initial evaluation of an electronic symptom diary for adolescents with cancer. JMIR Res Protoc 2012;1(2):e23 [FREE Full text] [doi: 10.2196/resprot.2175] [Medline: 23612521]

21. McMahon E, Wilson-Pauwels L, Henry JL, Jenkinson J, Sutherland B, Brierley M. The Iconic Pain Assessment Tool: Facilitating the Translation of Pain Sensations and Improving Patient-Physician Dialogue. J Bio Communication 2008;34:E20-E24.
22. Lalloo C, Henry JL. Evaluation of the Iconic Pain Assessment Tool by a heterogeneous group of people in pain. Pain Res Manag 2011;16(1):13-18 [FREE Full text] [Medline: 21369536]
23. Lalloo C, Stinson JN, Hochman JR, Adachi JD, Henry JL. Adapting the Iconic Pain Assessment Tool Version 2 (IPAT2) for adults and adolescents with arthritis pain through usability testing and refinement of pain quality icons. Clin J Pain 2013 Mar;29(3):253-264. [doi: 10.1097/AJP.0b013e318250e655] [Medline: 22936076]
24. Lalloo C, Stinson JN, Brown SC, Campbell F, Isaac L, Henry JL. Pain-QuILT: Assessing clinical feasibilty of a web-based tool for the visual self-report of pain in an interdisciplinary pediatric chronic pain clinic. Clinical J Pain 2014 (forthcoming). [Medline: 24300217]
25. Morgan DL. Focus groups as qualitative research. Planning and research design for focus groups. Thousand Oaks, CA: Sage; 1997:1-80.
26. Melzack R, Torgerson WS. On the language of pain. Anesthesiology 1971 Jan;34(1):50-59 [FREE Full text] [Medline: 4924784]
27. Melzack R, Katz J. The McGill Pain Questionnaire: Appraisal and current status. In: Handbook of pain assessment (2nd ed.). New York: Guilford Press; 2001:35-52.
28. Daut RL, Cleeland CS, Flanery RC. Development of the Wisconsin Brief Pain Questionnaire to assess pain in cancer and other diseases. Pain 1983 Oct;17(2):197-210. [Medline: 6646795]
29. Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, IMMPACT. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain 2005 Jan;113(1-2):9-19. [doi: 10.1016/j.pain.2004.09.012] [Medline: 15621359]
30. HyperRESEARCH: qualitative analysis tool computer program, Version 2.8. Massachusetts: ResearchWare Inc; 2009. URL: http://www.researchware.com/products/hyperresearch.html [accessed 2014-05-08] [WebCite Cache ID 6PQ3UEqY2]
31. Neuendorf K. The content analysis guidebook. Thousand Oaks, CA: Sage Publications; 2002.
32. PASW Statistics computer program, Version 17. Chicago: SPSS Inc; 2009. URL: http://www.spss.com.hk/statistics/ [accessed 2014-05-08] [WebCite Cache ID 6PQ43GL80]
33. Norman GR, Streiner D. Simple regression and correlation. In: Biostatistics: The Bare Essentials. Hamilton, ON: BC Decker Inc; 2000.
34. Lalloo C, Stinson JN, Hochman JR, Adachi JD, Henry JL. Adapting the Iconic Pain Assessment Tool Version 2 (IPAT2) for adults and adolescents with arthritis pain through usability testing and refinement of pain quality icons. Clin J Pain 2013 Mar;29(3):253-264. [doi: 10.1097/AJP.0b013e318250e655] [Medline: 22936076]
35. Feeley N, Cossette S, Côté J, Héon M, Stremler R, Martorella G, et al. The importance of piloting an RCT intervention. Can J Nurs Res 2009 Jun;41(2):85-99. [Medline: 19650515]
36. Wilkie DJ, Savedra MC, Holzemer WL, Tesler MD, Paul SM. Use of the McGill Pain Questionnaire to measure pain: a meta-analysis. Nurs Res 1990;39(1):36-41. [Medline: 2136771]
37. Carnes D, Parsons S, Ashby D, Breen A, Foster NE, Pincus T, et al. Chronic musculoskeletal pain rarely presents in a single body site: results from a UK population study. Rheumatology (Oxford) 2007 Jul;46(7):1168-1170 [FREE Full text] [doi: 10.1093/rheumatology/kem118] [Medline: 17488750]
38. Kamaleri Y, Natvig B, Ihlebaek CM, Bruusgaard D. Localized or widespread musculoskeletal pain: does it matter? Pain 2008 Aug 15;138(1):41-46. [doi: 10.1016/j.pain.2007.11.002] [Medline: 18077092]
39. Nakamura I, Nishioka K, Usui C, Osada K, Ichibayashi H, Ishida M, et al. An Epidemiological Internet Survey of Fibromyalgia and Chronic Pain in Japan. Arthritis Care Res (Hoboken) 2014 Jan 8:- (forthcoming). [doi: 10.1002/acr.22277] [Medline: 24403219]
40. Dorflinger L, Kerns RD, Auerbach SM. Providers' roles in enhancing patients' adherence to pain self management. Transl Behav Med 2013 Mar;3(1):39-46 [FREE Full text] [doi: 10.1007/s13142-012-0158-z] [Medline: 24073159]
41. Butow P, Sharpe L. The impact of communication on adherence in pain management. Pain 2013 Dec;154 Suppl 1:S101-S107. [doi: 10.1016/j.pain.2013.07.048] [Medline: 23928026]
42. Fox S, Duggan M. Pew Internet Internet Project. 2013. Tracking for Health URL: http://www.pewinternet.org/2013/01/28/tracking-for-health/ [accessed 2014-05-08] [WebCite Cache ID 6PQ4wP8rI]
43. 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]
44. Lalloo C, Jibb LA, Agarwal A, Stinson JN. There's a pain app for that: Review of Patient-Targeted Smartphone Applications for Pain Management. Clin J Pain 2014:- (forthcoming).
45. Peng P, Choiniere M, Dion D, Intrater H, Lefort S, Lynch M, STOPPAIN Investigators Group. Challenges in accessing multidisciplinary pain treatment facilities in Canada. Can J Anaesth 2007 Dec;54(12):977-984. [doi: 10.1007/BF03016631] [Medline: 18056206]
46. McMaster Industry Liaison Office (MILO). URL: http://ip.mcmaster.ca/ [accessed 2014-01-30] [WebCite Cache ID 6N0Aio54k]

47. Pain-QuILT Website. 2013. URL: http://painquilt.mcmaster.ca/ [accessed 2014-01-30] [WebCite Cache ID 6N0BFpAmg]

Abbreviations

BPI: Brief Pain Inventory
MPQ: McGill Pain Questionnaire
NRS: Numerical Rating Scale
SD: standard deviation

Edited by G Eysenbach; submitted 31.01.14; peer-reviewed by C von Baeyer, J Kelly; comments to author 15.03.14; revised version received 21.03.14; accepted 24.03.14; published 12.05.14.

Please cite as:
Lalloo C, Kumbhare D, Stinson JN, Henry JL
Pain-QuILT: Clinical Feasibility of a Web-Based Visual Pain Assessment Tool in Adults With Chronic Pain
J Med Internet Res 2014;16(5):e127
URL: http://www.jmir.org/2014/5/e127/
doi:10.2196/jmir.3292
PMID:24819478

Original Paper

Anonymity Versus Privacy: Selective Information Sharing in Online Cancer Communities

Jeana Frost[1], PhD; Ivar E Vermeulen[1], PhD; Nienke Beekers[2], MSc

[1]VU University Amsterdam, Amsterdam, Netherlands

[2]National IT Institute for Healthcare, Amsterdam, Netherlands

Corresponding Author:
Jeana Frost, PhD
VU University Amsterdam
De Boelelaan 1081
Amsterdam, 1081 HV
Netherlands
Phone: 31 20 5982782
Fax: 31 20 5983733
Email: j.h.frost@vu.nl

Abstract

Background: Active sharing in online cancer communities benefits patients. However, many patients refrain from sharing health information online due to privacy concerns. Existing research on privacy emphasizes data security and confidentiality, largely focusing on electronic medical records. Patient preferences around information sharing in online communities remain poorly understood. Consistent with the privacy calculus perspective adopted from e-commerce research, we suggest that patients approach online information sharing instrumentally, weighing privacy costs against participation benefits when deciding whether to share certain information. Consequently, we argue that patients prefer sharing clinical information over daily life and identity information that potentially compromises anonymity. Furthermore, we explore whether patients' prior experiences, age, health, and gender affect perceived privacy costs and thus willingness to share information.

Objective: The goal of the present study is to document patient preferences for sharing information within online health platforms.

Methods: A total of 115 cancer patients reported sharing intentions for 15 different types of information, demographics, health status, prior privacy experiences, expected community utility, and privacy concerns.

Results: Factor analysis on the 15 information types revealed 3 factors coinciding with 3 proposed information categories: clinical, daily life, and identity information. A within-subject ANOVA showed a strong preference for sharing clinical information compared to daily life and identity information ($F_{1,114}$=135.59, P=.001, η^2=.93). Also, adverse online privacy experiences, age, and health status negatively affected information-sharing intentions. Female patients shared information less willingly.

Conclusions: Respondents' information-sharing intentions depend on dispositional and situational factors. Patients share medical details more willingly than daily life or identity information. The results suggest the need to focus on anonymity rather than privacy in online communities.

(J Med Internet Res 2014;16(5):e126) doi:10.2196/jmir.2684

KEYWORDS

online systems; cancer; privacy; confidentiality; Health 2.0; anonymity

Introduction

Overview

Sharing information through patient platforms offers new opportunities for patients to learn about and manage their condition. However, information sharing online also introduces risks to patient privacy. There is a growing interdisciplinary scholarship around privacy that seeks to define the concept and construct systems to securely move data through a network. But, to date, there is a lack of research on user preferences concerning privacy and information sharing [1]. Building on previous work in the area of information systems and sharing [2-4], we propose and test a model for how patients think about

sharing information online. We suggest that patients will be most likely to share information when the benefits of doing so outweigh the risks. This cost benefit analysis is dynamic and varies according to who is sharing and the context of the exchange. Therefore, we argue that sharing preferences are best thought of in relational terms as a product of 3 factors: individual characteristics, type of information, and the breadth of the audience.

In this study, we look at the case of cancer patients and we distinguish 3 types of information: clinical, daily life, and identity information. We argue that although clinical information is "sensitive," it is most relevant to online health discussions and, therefore, it is shared most easily. Daily life and identity information are often shared in face-to-face conversations, but impose greater risk to anonymity and are less relevant for online health discussions. Therefore, they are shared less broadly. Individual factors further impact willingness to share both generally and around specific information types.

Background

Cancer rates are increasing across the developed world. In America, the lifetime risk of developing cancer is now 1 in 2 for men and 1 in 3 for women [5]. As the population in America and Europe ages, the cancer burden is expected to increase [6].

Cancer patients have a set of unmet psychosocial and informational needs that change over the course of treatment (for a review [7]). Although treatment needs near the time of diagnosis are often met, long-term physical, psychological, and psychosocial problems are given less attention [8]. The most frequently reported unmet needs of patients include psychological support [9,10], managing practical problems related to daily living, fear of recurrence [9], and information about genetics and the disease itself [11]. Studies also suggest the importance of addressing quality of life issues, both for improving survival rates (eg, [12]) and as a positive outcome measure [13,14]. For many of these unmet needs, there is no clear course of treatment or a "quick fix." Rather, patients and their families must learn to adapt, cope, and manage the variety of issues that arise.

Websites and Health 2.0 platforms are well-positioned to address these unmet patient needs. A growing set of engaged people, or ePatients, go online to search for information and connect with one another by exchanging health information and social support [15,16] on a variety of platforms (see Figure 1). Although findings are still mixed [17], a nascent body of literature suggests that these exchanges help address the otherwise unmet needs of cancer patients. In general, peer-to-peer support provides informational, emotional, and instrumental benefits [18]. For cancer patients specifically, online communities help in a variety of ways, including providing information on treatment and how to communicate with physicians, and emotional support on how to cope with cancer [19,20]. Despite some concern on the part of physicians, involvement in online communities seems to complement rather than replace the information and support coming from professionals. Indeed, research suggests that peers provide qualitatively different information [21] and support [22] than medical experts. Online communities are available anywhere, anytime, and potentially provide a place where sensitive topics can be safely discussed. In fact, people with sensitive problems who might have difficulty discussing these issues face-to-face are more likely to participate in online communities than people with conditions that are not stigmatized [23].

Active participation on online platforms appears to benefit both individual members and the community at large. Although passive viewing, or lurking, in an online community appears to be helpful for cancer patients [24], active participation in online communities has been linked to positive outcomes both online and offline, including improved mood [25], greater perceived online support, and offline improvements [26]. Perhaps more importantly, the value of a platform is tied to the level of user-generated content. Through sharing information and insight online, active participants improve the quality of the community. Therefore, it is important to promote active information sharing for the good of the individual members and the community at large.

Although Health 2.0 platforms present opportunities for patients, they also introduce the possibility of privacy invasions that could result in prejudice, decreases in economic opportunity, and potentially a loss of health care coverage. Past research suggests that concerns about privacy translate into online behavior—privacy concerns remain a key barrier to sharing information in online communities (eg, [27-29]). And privacy is a primary reason people cite for simply collecting information rather than actively participating online [30].

Despite the importance of privacy within online patient platforms and the Internet in general, users' preferences around information sharing are not well understood. Although online privacy is a rich interdisciplinary area of research, a literature review reveals a lack of attention to users' perspective in the design of privacy tools [1]. A similar gap exists in health care. Health care privacy research tends to treat privacy as a single construct, with an emphasis on protecting patient information to facilitate online exchanges. Most health care privacy research focuses implicitly or explicitly on data security within clinical systems, such as online electronic medical records and personal health records for which there are both moral and legal obligations to guard users from unintended harm (for a review [31]). The existing research on patient perspectives on privacy focuses on individual differences in willingness to share information [27,32], thereby, implicitly treating all types of information equally.

Legal and information systems research suggests that patient preferences might be complex. Legal scholars, noting an oversimplified use of the word "privacy" from the individual's perspective, have called for a more nuanced view—one that acknowledges the variety of types of invasions to privacy and examines the significance of privacy within a particular situation [33]. For the purpose of this study, we use Westin's concept of privacy: the right to privacy is the individual's ability to determine when, how, and to what extent information can be shared [34]. Related to the current work, research on consumers' willingness to disclose personal information on retail sites depends upon the nature of the information [35]. Building upon this legal and privacy scholarship, we propose that preferences

around sharing information in online health platforms are not uniform; rather, they vary with types of information and individual factors.

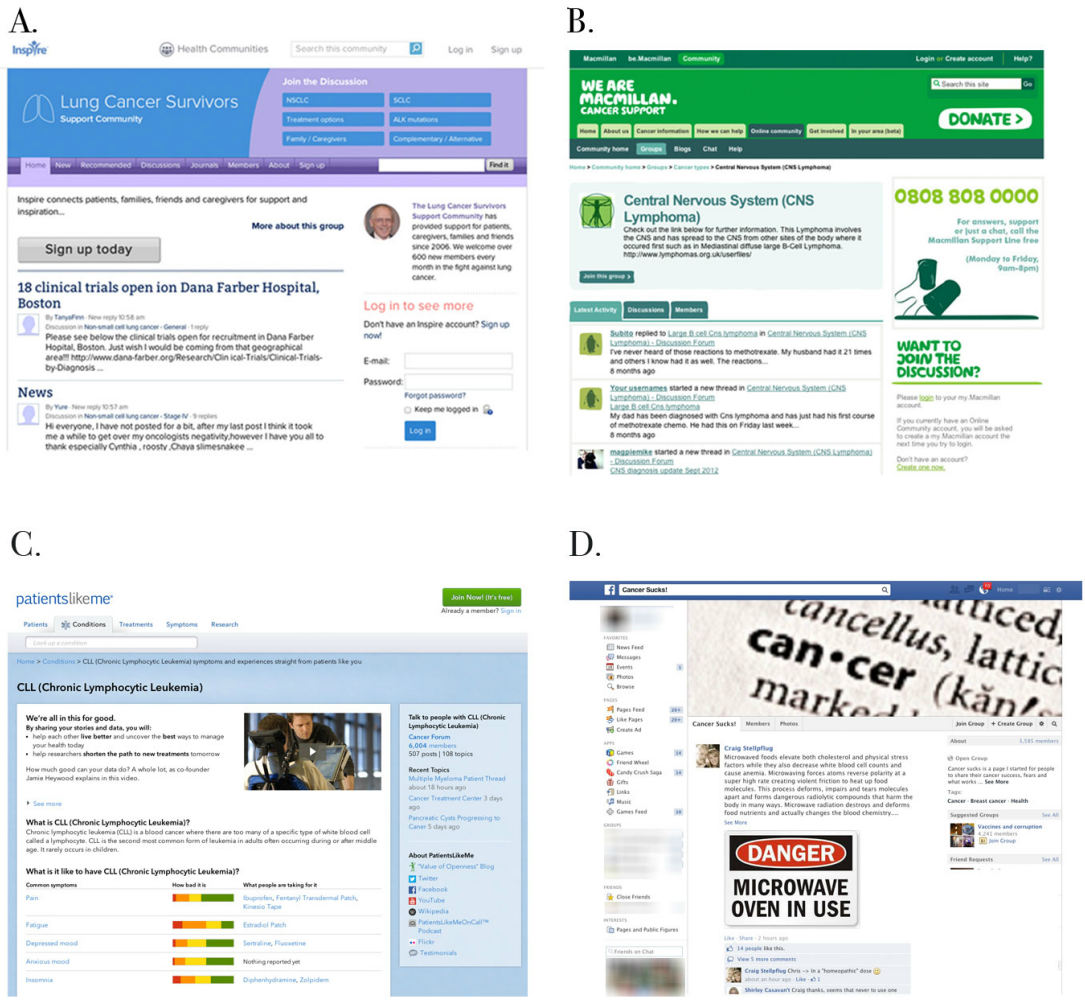
In particular, preferences may differ depending upon the perceived benefit of sharing a particular type of information with the anticipated audience. Within the context of e-commerce, people seem to conduct a mental calculation weighing the benefits against the costs of disclosing personal information to the system before making a purchase; this mental calculation has been labeled the "privacy calculus" [2]. In the health care domain, research suggests that patients choose to share information in situations when the expected value of sharing outweighs the possible risks [29].

In the current study, we will focus on 3 types of information sharing: clinical, daily life, and identity information. By clinical information, we mean detailed medical data describing diagnosis history, treatments, symptoms, and outcomes (eg, diagnosis data, cancer type, treatment regimen). We suggest that clinical information is pertinent to medical discussions and provides benefits for the online experience. When clinical information is shared without disclosing identity information (described subsequently), clinical information imposes low privacy costs. By daily life information, we mean information about professional life and relationships (eg, marital status and occupation). Although people may routinely share such information in casual face-to-face conversations, daily life information has marginal benefits for health conversations while introducing intermediate privacy costs. Finally, by identity information we mean information (eg, photo or personal email address) with little relevance in discussions about patient knowledge and psychological well-being, yet imposing high privacy costs—especially in combination with disclosed clinical information—by compromising user anonymity. Therefore, when viewed from the calculation of users weighing psychosocial and medical benefits against privacy costs, we expect that although some pieces of information may be relevant to more than one category, clinical information will be most easily shared, followed by daily life information, and then identity information.

In addition, we explore how personal characteristics and previous experiences online affect intentions around information sharing. First, intentions to share information may positively relate to the value patients anticipate from using a particular system and negatively relate to individuals' privacy concerns. Second, patients who have the highest expectation of life after cancer (eg, patients who are younger and have a better prognosis) may be the most reluctant to share health information [36]. Third, several studies suggest that women perceive higher online privacy costs than men [3,37,38], also in the context of health-related information [39]. Consequently, women may be less willing to share identity information than men.

In this study, we test patient preferences around privacy and anonymity with the central argument that people will be more interested in sharing information that is the stated topic of the online community. Furthermore, we expect that the willingness to share information depends on a combination of dispositional and situational factors [32]. To evaluate these hypotheses, we survey cancer patients interested in joining an online cancer community. Patients report on demographics, health status, expected utility of the forum, general privacy concerns, as well as on their willingness to share different types of information with different size audiences.

Figure 1. Screenshots of various Web platforms available for peer-to-peer discussion of cancer, including general platforms for many types of conditions including cancer (A and C), platforms for cancer specifically (B), and social media (D).



Methods

Participants

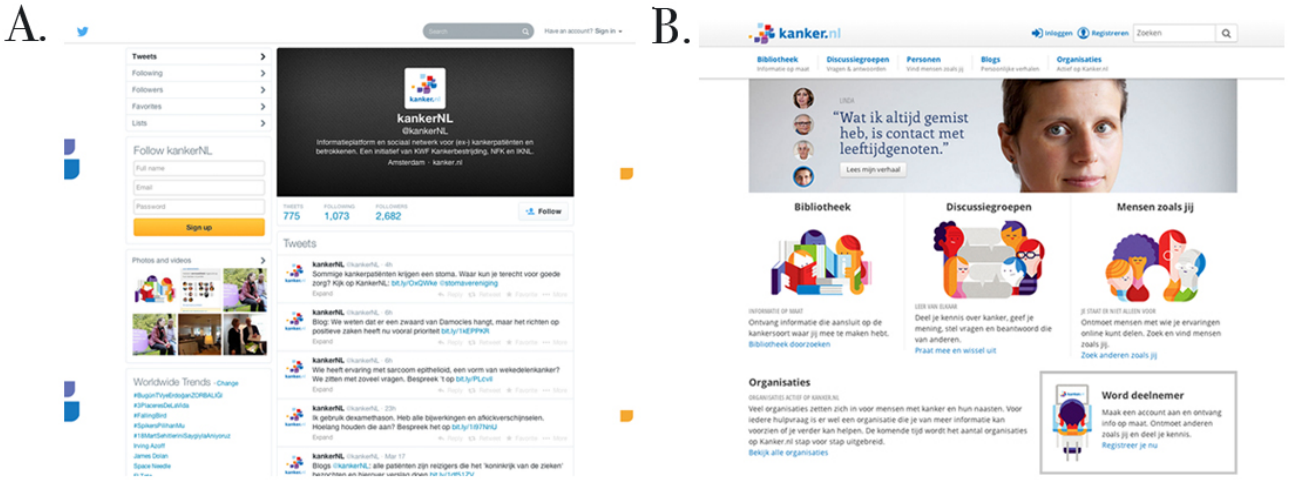
We recruited respondents through a website created to inform the design of a Dutch patient platform for cancer patients, Kanker.nl [40], and the platform's Twitter feed (see Figure 2). In total, 132 people completed the survey; 17 nonpatients were excluded from the analysis, leaving 115 current or previous cancer patients in the sample. The survey included the measures reported subsequently as well as items on the desired features of the Kanker.nl platform.

Table 1 displays the sample's characteristics. Participants had been diagnosed with a variety of cancer types, but 3 cancer groups were most prevalent: leukemia, bone marrow, and lymphoma (39.1%, 45/115); breast cancer (17.4%, 20/115); and cancers that affect the digestive organs (9.6%, 11/115). On average, the patients in the sample were diagnosed 6 (SD 6) years before the survey was conducted in 2006. Of all participants, 40.9% (47/115) were currently in treatment, 56.5% (65/115) had completed treatment, and 1.7% (2/115) could no longer be treated. The sample included more women (54.8%, 63/115) than men, and ranged in age (mean 52, SD 12, range 20-75).

Table 1. Baseline characteristics of the respondents (N=115).

Demographic variable	Response
Sex, n (%)	
Women	63 (54.8)
Men	52 (45.2)
Age (years), mean (SD)	52 (12)
Years since diagnosis, mean (SD)	6 (6)
Most commonly reported cancer types, n (%)	
Leukemia, bone marrow, and lymphoma	45 (39.1)
Breast cancer	20 (17.4)
Cancers that affected the digestive organs and	11 (9.6)
Treatment status, n (%)	
In treatment	47 (40.9)
Within 1 year of treatment	21 (18.2)
With 5 years of treatment	23 (20.0)
More than 5 years post treatment	21 (18.3)
Not treatable	2 (1.7)

Figure 2. Screenshots of the Kanker.nl Twitter feed (A) and the subsequently released Kanker.nl home page (B).



Instrument Development

The study employs a combination of existing and self-developed items (see Table 2). All items were translated into Dutch and entered into a Likert scale of 1 to 7, unless otherwise specified. Based on previous research demonstrating the reproducibility, reliability, and performance of a single item health status measure [28], we posed a single question asking respondents to rate their health status ("In general my health is..." with 1=very poor to 10=very good). The perceived usefulness of the future community was measured by using a question adapted from previous work ("How useful do you expect Kanker.nl will be for you?" with responses ranging from not at all to very useful) [29,21]. Health information privacy concerns were measured using 2 items (alpha=.84): (1) "I believe that submitting health information on the Internet is..." with responses ranging from not advisable at all to highly advisable and (2) "Health information on the Internet, once submitted..." with responses ranging from will not be misused at all to will be misused for sure [29]. Prior negative experiences with information sharing were measured using 1 item adapted from work on trust and information sources ("When it comes to the privacy invasion of health information, my online experience could be characterized as..." with responses ranging from no bad experiences to very bad experiences) [41]. Finally, patients reported on their phase of treatment by choosing between a closed set of options.

Table 2. Sources and text for all items except demographic variables and treatment phase (translated from Dutch).

Construct	Source	Item(s)	Response options
Health status	[42]	Overall, my health status is...	Poor to very good
Expected utility of the platform	[29,41]	How useful do you think kanker.nl will be for you?	Not at all to very useful
Prior negative experience	[23]	My experiences with privacy infringement on personal health can be described as...	Not at all negative to very negative
Privacy concern	[23]	I believe that submitting health information on the Internet is...	Highly advisable/not advisable at all
		Health information on the Internet, once submitted...	Will not be abused at all/will be abused for sure
Intention to share information	Self-developed	We are currently designing the privacy settings for Kanker.nl, for each piece of information, please indicate which group you would like to share it with	Only with personal contacts/only with members of the website/with all website visitors

To assess intentions to share different types of information, we asked participants about the breadth of the audience with whom they would like to share. To approximate the real world setting being studied, these questions were asked as if the respondents were choosing privacy settings for an online community. This community was described as a platform providing both expert and patient-generated information as well as peer-to-peer communication about cancer. Items were piloted in focus groups for clarity [27]. The response categories ranged from a smaller to larger audience, including (1) personal contacts, (2) members of the site, or (3) all website visitors. Participants were asked for sharing intentions with respect to 15 types of information: sex, age, marital status, family situation, profession, place of residence, province, picture, email address, patient status (ie, labeled as patient), type of cancer, date of diagnosis, treatment status, hospital, and clinician.

Using SPSS version 20 (IBM Corp, Armonk, NY, USA), we conducted a varimax-rotated principal component analysis (PCA), a variable reduction technique similar to factor analysis, to cluster together highly correlating items out of the 15 assessed. The PCA produced 3 factors with eigenvalues >1 together accounting for 73.05% of the variance (factor 1: 50.48%; factor 2: 13.34%; factor 3: 9.24%). Items were retained that loaded higher than 0.3 on their primary factor and had a primary loading of at least 0.2 higher than any of their cross-loadings.

Items pertaining to sharing clinical information (patient status, type of cancer, date of diagnosis, treatment status, sex, and age) loaded strongly on the first factor and were clustered together in a scale (alpha=.94). Items relating to sharing information about daily life (marital status, family situation, profession, and county) loaded strongly on the second factor (scale: alpha=.90). Items pertaining to sharing identity information (place of residence, picture, email address) loaded strongly on the third factor (scale: alpha=.78). Two items pertaining to sharing information about specific hospitals and clinicians loaded similarly on both the clinical and the identity dimension; thus, they were not included in either scale. In addition to the 3 subscales, we constructed a scale for overall sharing intentions of all 15 items (alpha=.93).

Results

A within-subject ANOVA showed that, consistent with our expectations, clinical information was most broadly shared (mean 2.32, SD 0.60, scale range 1-3), followed by daily life information (mean 1.86, SD 0.66), and identity information (mean 1.58, SD 0.56). The overall difference in intentions to share the 3 types of information was significant ($F_{1,114}$=135.59, P<.001, η^2=.93), as were the pairwise differences (all Ps<.001). See Table 3 for item-based responses.

Table 3. Willingness to share items: frequencies, means, and standard deviations.

Type of information	Willingness to share, n (%)			Mean (SD)
	Only with personal contacts	Only with members of the website	With all website visitors	
Clinical information				2.32 (0.60)
Patient status	11 (9.6)	49 (42.6)	55 (47.8)	2.38 (0.66)
Type of cancer	9 (7.8)	44 (38.3)	62 (53.9)	2.46 (0.64)
Date of diagnosis	15 (13.0)	61 (53.0)	39 (33.9)	2.21 (0.66)
Treatment status	17 (14.8)	58 (50.4)	40 (34.8)	2.20 (0.68)
Sex	12 (10.4)	44 (38.3)	59 (51.3)	2.41 (0.67)
Age	21 (18.3)	44 (38.3)	50 (43.5)	2.25 (0.75)
Daily life information				1.86 (0.67)
Marital status	50 (43.5)	38 (33.0)	27 (23.5)	1.80 (0.80)
Family situation	46 (40.0)	46 (40.0)	23 (20.0)	1.80 (0.75)
Profession	46 (40.0)	43 (37.4)	26 (22.6)	1.83 (0.78)
County	29 (25.2)	57 (49.6)	29 (25.2)	2.00 (0.71)
Identity information				1.58 (0.56)
Place of residence	52 (45.2)	49 (42.6)	14 (12.2)	1.67 (0.68)
Picture	65 (56.6)	36 (31.3)	14 (12.2)	1.56 (0.70)
Email address	66 (57.4)	40 (34.8)	9 (7.8)	1.50 (0.64)
Noncategorized information				
Specific hospital	20 (17.4)	63 (54.8)	32 (27.8)	2.10 (0.67)
Specific clinician	35 (30.4)	57 (49.6)	23 (20.0)	1.90 (0.70)

Table 4 shows the effects of individual differences on general sharing intentions. The model includes variables previously associated with differences in sharing preferences: expected value of the overall platform, general privacy concerns, health status, sex, age, and negative experiences online. Contrary to expectations, expected value of the platform, general privacy concerns, and sex (added as a dummy: male=–1, female=1) did not significantly predict general sharing intentions. However, results showed that prior negative experiences online (beta=–.43, P<.001) had a strong negative effect on sharing intentions. Also, older patients shared information more broadly than younger patients did (beta=.11, P=.01), and patients' health status had a marginal negative effect on general sharing intentions (beta=–.15, P=.07). The total regression model explained 25% (R=.54, adjusted R^2=.25) of the variance in intentions to share identity information ($F_{5,109}$=7.42, P<.001).

Table 4. Regression analysis: effects of interpersonal differences on cancer patients' general information-sharing intentions on an online platform.

Predictors	B	Beta	t_{113}	P
Expected utility of platform	.04	.08	0.99	.32
General privacy concerns	–.05	–.10	–1.14	.26
Negative experiences with online privacy	–.21	–.43	–4.97	<.001
Age	.11	.22	2.63	.01
Health status	–.07	–.15	–1.81	.07
Sex (dummy)	.30	.28	2.56	.01

Table 5 focuses on the effects of individual differences on intentions to share identity information. Contrary to expectations, the expected value of the platform did not produce the predicted positive effect. However, the effect of general privacy concerns (beta=–.19, P=.047) and adverse experiences with online privacy (beta=–.24, P=.008) negatively impacted sharing intentions as we expected. Also, older patients (beta=.24, P=.008) and patients with poorer health status (beta=–.26, P=.003) had fewer problems disclosing identity information, as did men (beta=–.19, P=.03). The total regression model explained 20% (R=.49, adjusted R^2=.20) of the variance in intentions to share identity information ($F_{5,109}$= 5.67, P<.001).

Finally, an exploratory *t* test showed no differences in sharing intentions for any information category between patients who were either pretreatment or during treatment versus posttreatment (all *t*s<1.24).

Table 5. Regression analysis: effects of interpersonal differences on cancer patients' intentions to share identity information on an online platform.

Predictors	B	Beta	t_{113}	*P*
Expected utility of platform	.01	.02	0.25	.81
General privacy concerns	–.08	–.19	–2.01	.047
Negative experiences with online privacy	–.14	–.24	–2.68	.008
Age	.14	.24	2.72	.008
Health status	–.15	–.26	–3.03	.003
Sex (dummy)	–.11	.19	–2.17	.03

Discussion

Principal Findings

This work suggests that privacy concerns are not uniform and depend on both individual and contextual factors. The privacy calculus model argues that people share information if the perceived rewards outweigh the perceived risks; as a result, people are selectively willing to share different types of information [2-4]. Although the selectivity in information sharing may be intuitive, the pattern of results is not. In the domain of patient platforms, the mental calculation about costs and benefits differs from other domains of life. During face-to-face conversations and in the context of generic online social platforms, such as Facebook, people routinely exchange names, ages, marital statuses, and professions. Yet, in the online health environment, ePatients are more willing to share what is often thought to be more sensitive health information. Platform members log on to share expertise based on medical experiences—details about their medical history and treatment are central to the experience. Indeed, although information may fit into more than 1 of the 3 identified domains, this study finds that patients are more willing to share clinical information than other forms of demographic and daily life information. Although patients mentioned in interviews that they would like to share daily life experiences with peers [36], the benefits of actively sharing such information with others appears lower than the benefits gained from sharing clinical information. As such, the privacy costs of sharing daily life information may often outweigh the possible benefits. In the case of identity information, risks of losing anonymity exceed the possible gains in the user experience and willingness to share is correspondingly low.

Our findings are consistent with previous research on the impact of prior experiences on willingness to share information [29] and sex [28]. Prior negative experiences regarding online privacy have a strong negative effect on patients' intentions to share all different types of information. Female respondents were less willing to share identity information than their male counterparts were. In general, women comprise the majority of social network users and tend to actively participate in online communities at a greater rate than men [43,44]. Yet, in this study and in previous studies, women seem to be more concerned with being identified than men. As a result, women are less willing to share information that potentially compromises anonymity [28].

An interesting finding is that patients who may be more concerned with their "life after cancer" (ie, younger patients and patients with a better health status) are less willing to share information with peers. This finding suggests that patients are fairly pragmatic in assessing the risks involved with disclosing information online. Although younger people tend to seek more information when making health-related decisions [45] and, therefore, may benefit more from an online community, the risks that sharing information imposes on future opportunities and possibly higher Internet literacy—including knowledge about privacy-related issues—may prevent younger and healthier patients from doing so.

Our results diverge from previous findings on the relationship between the expected value of the online platform on sharing intentions. This outcome suggests that patients do not associate sharing personal information with general site benefits, such as receiving more tailored peer feedback on their specific condition or situation. Instead, when contemplating which information to share with whom, patients may think about the value of that information for specific exchanges or the benefit of sharing that information for peers; that is, they could be considering to what extent specific types of information may benefit a particular online conversation or other community members. Also, our results showed that general privacy concerns hamper sharing intentions only with respect to the specific category of identity information and not with respect to general sharing intentions. This outcome further supports the notion that all information sharing is not equal.

Limitations

In this study, we examined intentions and preferences around sharing information within an online patient community through a survey. A basic limitation of our approach is the hypothetical nature of the questions asked. Respondents were recruited from a group of people interested in joining an online patient platform and the questions about privacy resembled those asked in an online community. However, our survey was conducted outside the context of such an online community, and asked for potential users' sharing intentions rather than measuring actual sharing behaviors. In other settings, people routinely deviate from stated privacy preferences in their actual behavior [46]. Although patients may perform a cost benefit analysis when they are

thoughtfully contemplating privacy settings in a survey, people may behave differently in actual online communities.

A second limitation to the study concerns the sampling of respondents. The respondents were members of an online panel of people interested in joining a new patient platform and volunteered for the study. As such, these particular individuals may have a higher level of Internet literacy than the patient population at large. Still, this sample seems representative of the population likely to participate in current and emerging online communities; indeed, most of our respondents expressed a desire to participate in the Kanker.nl platform.

Future work could examine the role of additional factors that influence sharing intentions and study sharing behavior directly. Emotional drivers (eg, fear, shame) may impact willingness to share within a peer-to-peer environment [47], complementing the more rational cost-benefit approach presented in the current research. Prosocial behavior and willingness to disclose information is associated with peers who do the same [48]. Through this process of disclosure, norms of trust may emerge. Consequently, future research could also look into reciprocity as a driver of willingness to share.

Conclusions

The current study extends our understanding of patient privacy preferences by disaggregating the notion of privacy concerns. Different types of information pose different concerns, and different users have different concerns. Because of continual changes in technology and the user base of online health communities, understanding privacy concerns and sharing behaviors from a patient perspective is an ongoing process. However, such understanding is crucial to optimize the positive effects online health communities have on well-being.

This work has implications for designing online communities. Research shows that in many patient populations (eg, posttreatment patients) there are clear benefits to reading and reviewing content from online communities [24]. Several types of Health 2.0 platforms and communities now exist, some of which protect identity to such an extent that patients can only log on individually to access medical records, maintaining confidentiality (eg, Kaiser Permanente's My Health Manager and the Mayo Clinic's Patient Online Services), whereas others piggyback on existing social networks with no pretense of anonymity (eg, Genentech's Circle of Support App for breast cancer patients and the Cancer Sucks! community on Facebook). The current research suggests an intermediate solution. To align with patient preferences, systems should guard anonymity while facilitating clinical information sharing.

Conflicts of Interest

None declared.

References

1. Pavlou PA. State of the information privacy literature: where are we now and where should we go. MIS Quarterly 2011;35(4):977-988.
2. Dinev T, Hart P. An extended privacy calculus model for e-commerce transactions. Information Systems Research 2006;17(1):A.
3. Dinev T, Hart P. Internet privacy concerns and their antecedents - measurement validity and a regression model. Behaviour & Information Technology 2004 Nov;23(6):413-422. [doi: 10.1080/01449290410001715723]
4. Dinev T, Bellotto M, Hart P, Russo V, Serra I, Colautti C. Privacy calculus model in e-commerce – a study of Italy and the United States. Eur J Inf Syst 2006 Aug;15(4):389-402. [doi: 10.1057/palgrave.ejis.3000590]
5. Howlader N, Noone AM, Krapcho M, Neyman N, Aminou R, Waldron W, et al. Surveillance, Epidemiology, and End Results Program. Bethesda: National Cancer Institute; 2011. SEER cancer statistics review URL: http://seer.cancer.gov/archive/csr/1975_2008/ [accessed 2014-04-04] [WebCite Cache ID 6Oa1q2jTF]
6. Ferlay J, Autier P, Boniol M, Heanue M, Colombet M, Boyle P. Estimates of the cancer incidence and mortality in Europe in 2006. Ann Oncol 2007 Mar;18(3):581-592 [FREE Full text] [doi: 10.1093/annonc/mdl498] [Medline: 17287242]
7. Rutten LJ, Arora NK, Bakos AD, Aziz N, Rowland J. Information needs and sources of information among cancer patients: a systematic review of research (1980-2003). Patient Educ Couns 2005 Jun;57(3):250-261. [doi: 10.1016/j.pec.2004.06.006] [Medline: 15893206]
8. Stanton AL. Psychosocial concerns and interventions for cancer survivors. J Clin Oncol 2006 Nov 10;24(32):5132-5137. [doi: 10.1200/JCO.2006.06.8775] [Medline: 17093275]
9. Armes J, Crowe M, Colbourne L, Morgan H, Murrells T, Oakley C, et al. Patients' supportive care needs beyond the end of cancer treatment: a prospective, longitudinal survey. J Clin Oncol 2009 Dec 20;27(36):6172-6179. [doi: 10.1200/JCO.2009.22.5151] [Medline: 19884548]
10. Sanson-Fisher R, Girgis A, Boyes A, Bonevski B, Burton L, Cook P. The unmet supportive care needs of patients with cancer. Supportive Care Review Group. Cancer 2000 Jan 1;88(1):226-237. [Medline: 10618627]
11. Harrison JD, Young JM, Price MA, Butow PN, Solomon MJ. What are the unmet supportive care needs of people with cancer? A systematic review. Support Care Cancer 2009 Aug;17(8):1117-1128. [doi: 10.1007/s00520-009-0615-5] [Medline: 19319577]

12. Ganz PA, Lee JJ, Siau J. Quality of life assessment. An independent prognostic variable for survival in lung cancer. Cancer 1991 Jun 15;67(12):3131-3135. [Medline: 1710541]
13. Demetri GD, Kris M, Wade J, Degos L, Cella D. Quality-of-life benefit in chemotherapy patients treated with epoetin alfa is independent of disease response or tumor type: results from a prospective community oncology study. Procrit Study Group. J Clin Oncol 1998 Oct;16(10):3412-3425. [Medline: 9779721]
14. Forsythe LP, Alfano CM, Leach CR, Ganz PA, Stefanek ME, Rowland JH. Who provides psychosocial follow-up care for post-treatment cancer survivors? A survey of medical oncologists and primary care physicians. J Clin Oncol 2012 Aug 10;30(23):2897-2905 [FREE Full text] [doi: 10.1200/JCO.2011.39.9832] [Medline: 22778322]
15. Ferguson T, Frydman G. The first generation of e-patients: These new medical colleagues could provide sustainable healthcare solutions. BMJ: British Medical Journal 2004;328(7449):1148.
16. 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]
17. Hoey LM, Ieropoli SC, White VM, Jefford M. Systematic review of peer-support programs for people with cancer. Patient Educ Couns 2008 Mar;70(3):315-337. [doi: 10.1016/j.pec.2007.11.016] [Medline: 18191527]
18. Campbell HS, Phaneuf MR, Deane K. Cancer peer support programs-do they work? Patient Educ Couns 2004 Oct;55(1):3-15. [doi: 10.1016/j.pec.2003.10.001] [Medline: 15476984]
19. Meier A, Lyons EJ, Frydman G, Forlenza M, Rimer BK. How cancer survivors provide support on cancer-related Internet mailing lists. J Med Internet Res 2007;9(2):e12 [FREE Full text] [doi: 10.2196/jmir.9.2.e12] [Medline: 17513283]
20. Klemm P, Bunnell D, Cullen M, Soneji R, Gibbons P, Holecek A. Online cancer support groups: a review of the research literature. Comput Inform Nurs 2003;21(3):136-142. [Medline: 12792194]
21. Hartzler A, Pratt W. Managing the personal side of health: how patient expertise differs from the expertise of clinicians. J Med Internet Res 2011;13(3):e62 [FREE Full text] [doi: 10.2196/jmir.1728] [Medline: 21846635]
22. Shaw BR, McTavish F, Hawkins R, Gustafson DH, Pingree S. Experiences of women with breast cancer: exchanging social support over the CHESS computer network. J Health Commun 2000;5(2):135-159. [doi: 10.1080/108107300406866] [Medline: 11010346]
23. Davison KP, Pennebaker JW, Dickerson SS. Who talks? The social psychology of illness support groups. Am Psychol 2000 Feb;55(2):205-217. [Medline: 10717968]
24. van Uden-Kraan CF, Drossaert CH, Taal E, Seydel ER, van de Laar MA. Self-reported differences in empowerment between lurkers and posters in online patient support groups. J Med Internet Res 2008 Jun;10(2):e18 [FREE Full text] [doi: 10.2196/jmir.992] [Medline: 18653442]
25. Rodgers S, Chen Q. Internet community group participation: Psychosocial benefits for women with breast cancer. Journal of Computer ‐ Mediated Communication 2005;10(4):00-00. [doi: 10.1111/j.1083-6101.2005.tb00268.x]
26. Nimrod G. Challenging the Internet paradox: Online depression communities and well-being. International Journal of Internet Science 2013;8(1):30-48.
27. Stewart KA, Segars AH. An empirical examination of the concern for information privacy instrument. Information Systems Research 2002 Mar;13(1):36-49. [doi: 10.1287/isre.13.1.36.97]
28. Cho H, Rivera-Sanchez M, Lim SS. A multinational study on online privacy: global concerns and local responses. New Media & Society 2009 Apr 28;11(3):395-416. [doi: 10.1177/1461444808101618]
29. Bansal G, Zahedi FM, Gefen D. The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online. Decision Support Systems 2010 May;49(2):138-150. [doi: 10.1016/j.dss.2010.01.010]
30. Nonnecke B, Preece J. Shedding light on lurkers in online communities. Ethnographic Studies in Real and Virtual Environments: Inhabited Information Spaces and Connected Communities, Edinburgh 1999:123-128.
31. Appari A, Johnson ME. Information security and privacy in healthcare: current state of research. IJIEM 2010;6(4):279-314. [doi: 10.1504/IJIEM.2010.035624]
32. Culnan MJ. How did they get my name?: An exploratory investigation of consumer attitudes toward secondary information use. MIS Quarterly 1993;17(3):341-363.
33. Solove DJ. A taxonomy of privacy. University of Pennsylvania Law Review 2006;154(3):477-564.
34. Westin AF. Privacy and freedom. Washington and Lee Law Review 1968;25(1):166.
35. Malhotra NK, Kim SS, Agarwal J. Internet users' information privacy concerns (IUIPC): the construct, the scale, and a causal model. Information Systems Research 2004;15(4):336-355.
36. Frost J, Beekers N, Hengst B, Vendeloo R. Meeting cancer patient needs: Designing a patient platform. 2012 Presented at: CHI '12 Extended Abstracts on Human Factors in Computing Systems; May 5-10, 2012; Austin, TX URL: http://dl.acm.org/citation.cfm?id=2223806 [doi: 10.1145/2212776.2223806]
37. Hoy, Milne G. Gender differences in privacy-related measures for young adult Facebook users. Journal of Interactive Advertising 2010 Mar;10(2):28-45. [doi: 10.1080/15252019.2010.10722168]
38. Fogel J, Nehmad E. Internet social network communities: Risk taking, trust, and privacy concerns. Computers in Human Behavior 2009 Jan;25(1):153-160. [doi: 10.1016/j.chb.2008.08.006]
39. Laric MV, Pitta D, Katsanis L. Consumer concerns for healthcare information privacy: A comparison of US and Canadian perspectives. Research in Healthcare Financial Management 2009;12(1):93-111.

40. Kanker.nl. URL: https://www.kanker.nl/ [accessed 2014-04-01] [WebCite Cache ID 6OVcwMZLw]
41. Song J, Zahedi F. Trust in health infomediaries. Decision Support Systems 2007;43(2):390-407.
42. DeSalvo KB, Fisher WP, Tran K, Bloser N, Merrill W, Peabody J. Assessing measurement properties of two single-item general health measures. Qual Life Res 2006 Mar;15(2):191-201. [doi: 10.1007/s11136-005-0887-2] [Medline: 16468076]
43. Hampton K, Goulet LS, Rainie L, Purcell K. Social networking sites and our lives. Washington, DC: Pew Internet & American Life Project; 2012 Jun 16. URL: http://www.pewinternet.org/2011/06/16/social-networking-sites-and-our-lives/ [accessed 2014-03-10] [WebCite Cache ID 6NyY1QJC3]
44. Kimbrough AM, Guadagno RE, Muscanell NL, Dill J. Gender differences in mediated communication: Women connect more than do men. Computers in Human Behavior 2013 May;29(3):896-900. [doi: 10.1016/j.chb.2012.12.005]
45. Carstensen LL, Lackenhoff CE. Aging, emotion, and evolution. Annals of the New York Academy of Sciences 2006;1000(1):152-179. [doi: 10.1196/annals.1280.008]
46. Norberg PA, Horne DR, Horne DA. The privacy paradox: Personal information disclosure intentions versus behaviors. Journal of Consumer Affairs 2007;41(1):100-126. [doi: 10.1111/j.1745-6606.2006.00070.x]
47. Anderson CL, Agarwal R. The digitization of healthcare: boundary risks, emotion, and consumer willingness to disclose personal health information. Information Systems Research 2011;22(3):469-490.
48. Chen YR, Chen XP, Portnoy R. To whom do positive norm and negative norm of reciprocity apply? Effects of inequitable offer, relationship, and relational-self orientation. Journal of Experimental Social Psychology 2009 Jan;45(1):24-34. [doi: 10.1016/j.jesp.2008.07.024]

Abbreviations

PCA: principal component analysis

Edited by G Eysenbach; submitted 22.04.13; peer-reviewed by E Gibson, J Gao, M Morrison; comments to author 15.06.13; revised version received 07.02.14; accepted 27.02.14; published 14.05.14.

Please cite as:
Frost J, Vermeulen IE, Beekers N
Anonymity Versus Privacy: Selective Information Sharing in Online Cancer Communities
J Med Internet Res 2014;16(5):e126
URL: http://www.jmir.org/2014/5/e126/
doi:10.2196/jmir.2684
PMID:24828114

Original Paper

Evaluating a Web-Based Clinical Decision Support System for Language Disorders Screening in a Nursery School

María Luisa Martín Ruiz[1]; Miguel Ángel Valero Duboy[1], PhD; Carmen Torcal Loriente[2]; Iván Pau de la Cruz[1], PhD

[1]Technical University of Madrid, Department of Telematic and Electronic Engineering, Madrid, Spain

[2]Legamar School, Madrid, Spain

Corresponding Author:
María Luisa Martín Ruiz
Technical University of Madrid
Department of Telematic and Electronic Engineering
Carretera de Valencia KM 7
Madrid, 28031
Spain
Phone: 34 91336 3770
Fax: 34 91336 7817
Email: marialuisa.martinr@upm.es

Abstract

Background: Early and effective identification of developmental disorders during childhood remains a critical task for the international community. The second highest prevalence of common developmental disorders in children are language delays, which are frequently the first symptoms of a possible disorder.

Objective: This paper evaluates a Web-based Clinical Decision Support System (CDSS) whose aim is to enhance the screening of language disorders at a nursery school. The common lack of early diagnosis of language disorders led us to deploy an easy-to-use CDSS in order to evaluate its accuracy in early detection of language pathologies. This CDSS can be used by pediatricians to support the screening of language disorders in primary care.

Methods: This paper details the evaluation results of the "Gades" CDSS at a nursery school with 146 children, 12 educators, and 1 language therapist. The methodology embraces two consecutive phases. The first stage involves the observation of each child's language abilities, carried out by the educators, to facilitate the evaluation of language acquisition level performed by a language therapist. Next, the same language therapist evaluates the reliability of the observed results.

Results: The Gades CDSS was integrated to provide the language therapist with the required clinical information. The validation process showed a global 83.6% (122/146) success rate in language evaluation and a 7% (7/94) rate of non-accepted system decisions within the range of children from 0 to 3 years old. The system helped language therapists to identify new children with potential disorders who required further evaluation. This process will revalidate the CDSS output and allow the enhancement of early detection of language disorders in children. The system does need minor refinement, since the therapists disagreed with some questions from the CDSS knowledge base (KB) and suggested adding a few questions about speech production and pragmatic abilities. The refinement of the KB will address these issues and include the requested improvements, with the support of the experts who took part in the original KB development.

Conclusions: This research demonstrated the benefit of a Web-based CDSS to monitor children's neurodevelopment via the early detection of language delays at a nursery school. Current next steps focus on the design of a model that includes pseudo auto-learning capacity, supervised by experts.

(J Med Internet Res 2014;16(5):e139) doi:10.2196/jmir.3263

KEYWORDS

primary health care; health information systems; knowledge management; evaluation; early diagnosis; eHealth; language disorders

Introduction

The early detection of neurodevelopmental disorders in childhood is a key task to support diagnosis and treatment processes [1,2]. The substantial role of language development, from zero until the age of 6 years, strongly influences communication and social skills in children and adults [3,4]. Furthermore, experience shows that language acquisition delays do influence social and behavioral attitudes, lack of school readiness [5], school exclusion [6], future academic problems [7], neuropsychiatric disorders [8], and poor employment [9]. Although diverse medical procedures aim to support the detection of neurological disorders in children [2,10-12], there is a lack of adoption at the primary pediatric care level, as it requires too much time and specialized knowledge [13].

Estimates of disability predominance in childhood vary due to differences in definition and the wide range of methodologies and existing measuring instruments [14]. The World Health Organization (WHO) and the World Bank declare, in their report, "World report on disability", that many cases of children with disability are still not identified and do not receive diagnosis or treatment services from health care entities [14]. Hence, early and effective identification of children with developmental disorders remains a critical task for the international community. Language disorders are frequently the first symptoms of a possible developmental disorder [15]. The prevalence of language delays is the second highest within common developmental disorders (1-19%) [16] and it is often associated with negative long-term outcomes [3,17-19].

The mental health, social, and behavioral developmental needs of very young children have gained awareness in the last 10 years [1,20,21]. Moreover, the acquisition of communication skills is essential for all students due to its direct impact on school success [22]. Thus, the early detection of developmental disorders in early childhood may facilitate the necessary diagnosis and/or treatment actions [13], as well as the early adoption of educational recommendations and activities for professionals and parents.

Most children achieve good verbal communication by the age of 3 years [3,23], although language acquisition level has a variable range within a target population. Hence, the availability of an effective language development CDSS may facilitate early identification of these types of disorders before the age of 3 years. Both primary care and education professionals can play a valuable role in early detection during their regular interactions with a child. Unfortunately, the lack of resources to perform individualized exhaustive evaluations of all children makes the use of efficient and reliable methods of detection necessary [4]. So far as this is concerned, diverse studies demonstrate that teachers can identify pupils with language difficulties, with sufficient precision and sensibility, if they have been provided with a guide or suitable orientation [24,25].

Ygual et al discovered a significant correlation between teachers' observations and criteria scores on intelligibility, literal understanding of sentences, grammatical expression, and lexical richness [4]. The research published by Wilson et al [3] reinforced the argument that early interventions can affect long-term outcomes, and concluded that language delay is not easily predictable from available risk factors. Therefore, it is not possible to foresee whether a child will have a language delay at 30 months and the identification of this disorder would require direct clinical contact with all families [3]. The evolution of the effectiveness of this kind of solution has been demonstrated for years, as initially explored by Miller 20 years ago [26], and surveyed by Berner and Maisiak in 1999 who concluded that a CDSS can function both to confirm and to broaden physicians' diagnostic thinking [27].

The main objective of this research was to evaluate the deployment of a Web-based system for efficient screening of language disorders at the early stages of a child's development. The implemented solution is a Clinical Decision Support System (CDSS), called "Gades", whose use was widely tested in a nursery school. This paper discusses the results obtained from the Gades validation to provide professionals with real-time knowledge on early identification of 146 children with possible language disorders. The previous publication of Gades' user requirements, implementation, deployment, and validation showed high success from a usability point of view [13].

The development of a knowledge base (KB) [28], needed to build the CDSS, relied on incremental interactions and refinement with the experts community. A set of 41 retrospective cases of children, treated over 15 years at the Language Intervention Center (LIC) of La Salle University, Madrid, helped to fine-tune the questions in the KB, starting with well-accepted neurodevelopmental tests.

Methods

Gades Knowledge Base

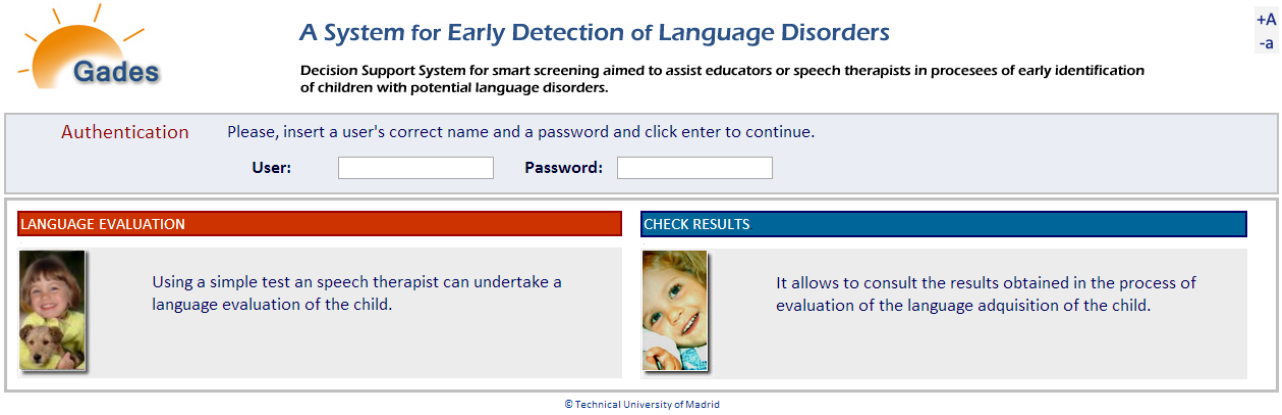
Gades KB is based on an ontology that integrates a child's language acquisition information according to age. It has over 100 rules to generate alerts and/or alarms in case deviations from the child's development stage are detected. The initial version of Gades KB was built between September 2011 and April 2013, according to the experience of a multidisciplinary team. Two language therapists, a neuropediatrician, a neonatologist, and three engineers supplied inputs for Gades KB, updating original versions of the Denver Developmental Screening Test. The team used CommonKADS (CK) methodology [29], to design and develop a decision support system based on the knowledge extracted from human experts and its required codification for system processing [29]. The baseline of Gades KB was Denver Test, as it is widely accepted in primary pediatric care [30]. The Gades KB takes advantage of the monthly structured questions of Denver. Each of the items of Denver represent the mean of the language development for each month of the child's life. The first version of the KB included 136 questions, from month 1 to month 72. These questions belong to two main categories: (1) questions called "Alert Milestones" that recommend a visit to the pediatrician. A negative answer to these developmental items means that the child should make a regular follow-up visit within 3 months to re-evaluate the level of language acquisition, and (2) questions called "Alarm Milestones", that suggest direct referral to a specialist in language disorders.

The Gades KB evaluates four areas of speech and language development: Sensory Reception (SR), Speech Perception (SP), Speech Production (SPD), and Pragmatic (P). The Gades system relies on Gades KB to support early detection of language disorders.

Gades Clinical Decision Support Web-Based System

The Gades system aims to enhance early detection of children with language disorders. Its evaluation process involved 146 children attending Legamar Nursery School. The Gades KB integrates all the knowledge and logic associated with the decisions supported by the system. The potential outcome is the suggestion of early referral to specialists if a child under 6 years old may have a language disorder. Figure 1 shows the Gades Web interface whose home page includes the following functionality: user authentication, language evaluation, and results obtainment.

Figure 1. Gades home page and main functionality.

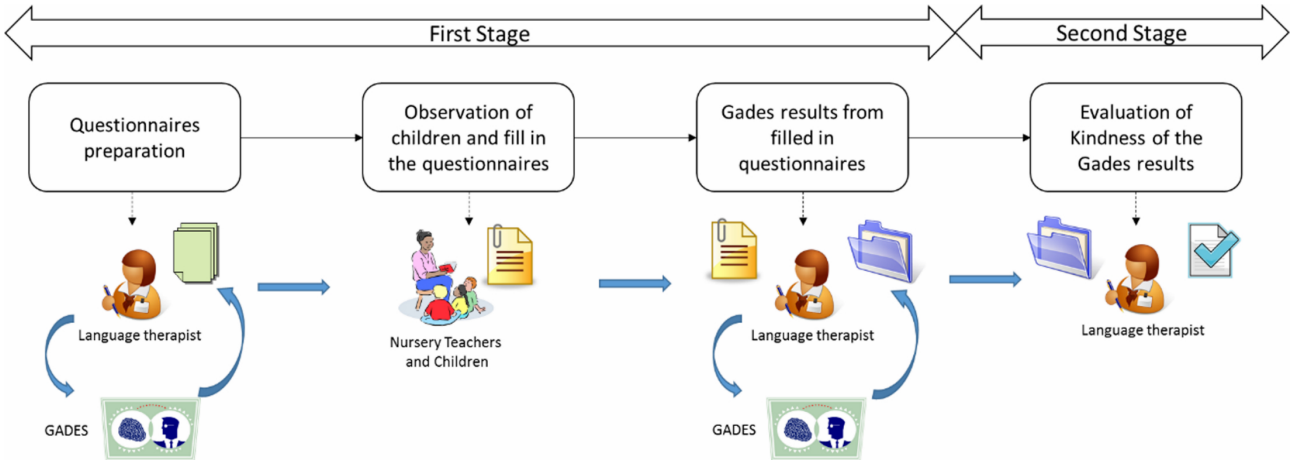


Gades Deployment and Evaluation Method

A Nursery School Language Therapist (NSLT), employed at Legamar Nursery School, Madrid, evaluated the Gades system in the spring of 2013. A total of 63 boys and 83 girls participated in the study; 94 children from 0 to 3 years old and 52 children in the 4-6 years stage. The number of enrolled children in the 0-3 years group was higher since the early detection of language disorders is a research priority in this developmental period. The entire staff of educators at Legamar observed and evaluated the behavior of the children, by following the questions suggested by Gades. The average age of the 12 teachers at Legamar was 34 years old. All the educators and the NSLT were women with little background in information technology.

The study started 6 months after the beginning of the school year to ensure that teachers had enough information about their pupils. Figure 2 shows the two stages of the methodology, with the same NSLT in both. The first stage involved language evaluation of all the children. The Gades KB helped the NSLT to obtain questionnaires for every group of children who participated in the study. Educators received the paper questionnaires along with an initial training session. The NSLT proposed child observation for one week before starting the evaluation process. After the observation period, educators filled out one questionnaire per child and gave them back to the NSLT. The questionnaires provided by the educators with their perceptions of the child enhanced the information acquisition process and gave the NSLT better evidence for each child enrolled in the study. The NLST updated the children's data in Gades to avoid usability problems or system interaction barriers. In the second stage, the accuracy of the Gades results was evaluated. The NSLT validated the results for each evaluation and stated whether or not she agreed with the Gades evaluation. The NSLT also checked the questions relating to the aforementioned results. When the NSLT did not agree with the Gades decisions, she analyzed the language areas evaluated by the KB questions. Thus, the NSLT considered non-evaluated language areas and proposed modifications in order to improve the Gades KB.

Figure 2. Evaluation method for early detection of language disorders.



Sex and Age of the Population Enrolled in the Study

Table 1 summarizes the population and number of language evaluations carried out by age and sex. All children at Legamar participated in the study and their distribution shows a higher ratio of girls at all ages (except the age of 5 years where there were more boys). Overall, 56.8% (83/146) of evaluated children were girls and 43.2% (63/146) were boys.

In terms of age, the number of subjects involved in the study at age 3 was higher (48 children from the sample of 146 subjects), since there was a higher number of children enrolled at nursery school at this age. From a prevention point of view, the detection of language disorders before the age of 3 years is a key issue as it directly influences the Quality of Life related to Health (QoLrH). Hence, 100% of the children between 0 and 3 years at Legamar participated in the research. However, only 37.1% (52/140) of the children between 4-6 years were included, as early detection is not significant at this stage. In the 0-3 years stage, a population of 94 children participated in the study, and the distribution of children who attended the nursery school was 59% (55/94) girls and 41% (39/94) boys. In the 4-6 years stage, 52 children of the nursery school received language evaluation by Gades: 54% (28/52) were girls and 46% (24/52) were boys.

Table 1. Number of language evaluations conducted by age and sex.

	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Total
Boys	5	14	20	7	14	3	63
Girls	6	21	28	14	8	6	83
Total	11	35	48	21	22	9	146

Results

Overview

The following section details the language evaluation results, obtained from the feedback supplied by the NSLT and teachers who participated in the research. Table 2 relates the global results extracted from the evaluation of 146 children with the Gades decision acceptance ratio, classified by developmental stages and years. The NLST accepted 122 cases resulting in an 84% success ratio. This ratio is even higher at the 0-3 years stage, where 87 cases (93%) were accepted from a total of 94. The NSLT disagreed with the Gades outcome in 24 cases (16%). The higher concentration of this non-acceptance ratio focuses on the 4-6 years stage (17/24). As previously noted, the NLST agreed with the decision in 87 cases at the 0-3 years stage and rejected Gades outcomes in 7 cases (7%). The successful cases reached a total of 35 children in the 4-6 years stage with 17 decisions requiring additional review. The total number of cases between years 5 and 6 is lower, with 9 evaluated children, since the experiment did not include children who had reached 6 years old before the spring term. The ratio of non-accepted decisions is too high at this stage (8/9), which suggests the need to improve the sample and KB for this group. The best Gades outcomes happened in the 25-36 months group where the NLST positively accepted all the suggestions (48). Table 2 provides further details with absolute and comparable relative data.

Although the NSLT agreed with all the questions for several months, she pointed out the need to refine the KB in order to improve Gades' decisions. For example, discrepancies arose between Gades' decision and the NSLT—the therapist recommended postponing some questions, which are not yet required for some months in the second year. At years 4, 5, and 6, the NSLT requested adding questions related to the articulation of the language and pragmatic ability. The KB refinement will require cooperation between the NSLT at Legamar and other speech therapists.

A key result obtained from the Gades evaluation was the identification of possible language delays in 7 children who had not previously caused alarm to either the NSLT or his/her educator. These cases require a formal diagnosis process in order to compare the system's decision with traditional methods. These children had been enrolled for a few months at Legamar and the speech therapist had not noticed any delays. Table 3

summarizes the decisions provided by the NSLT after the formal evaluation.

The NSLT identified discrepancies in cases 1, 2, and 3, described in Table 3, between the behavior observed in the school environment and the behavior confirmed by parents. A typical explanation is the deviation of the linguistic functionality of some children, mediatized by the difference between the language used at the school and the one used at home by the family. The reliability of the observation, carried out by parents or relatives, always needs to be checked to avoid subjective approaches. Gades' outcomes led to the initiation of early therapeutic actions at Legamar in cases 4 to 7.

Table 2. Number of language evaluations performed and therapist decisions.

Stage	Months of the questions	Year (age) of the child	Total number of cases	Number of cases where NSLT[a] accepted Gades decision (%)	Number of cases where NSLT did not accept Gades decision (%)
0-3 years	0-12	Year 1	11	11 (100.0)	0 (0.0)
	13-24	Year 2	35	28 (80.0)	7 (20.0)[b]
	25-36	Year 3	48	48 (100.0)	0 (0.0)
	Total		94	87 (92.6)	7 (7.4)
4-6 years	37-48	Year 4	21	13 (61.9)	8 (38.1)[c]
	49-60	Year 5	22	21 (95.5)	1 (4.5)[c]
	61-72	Year 6	9	1 (11.1)	8 (88.9)[c]
	Total		52	35 (67.3)	17 (32.7)
Total			146	122 (83.6)	24 (16.4)

[a]NLST: nursery school language therapist.

[b]The age in months of some questions is incorrect. Therefore, the NSLT believed that some questions should be delayed.

[c]Incorporating additional questions related to the articulation of language and pragmatics is required.

Table 3. Decision of the NSLT[a] on 7 new cases of children with possible language delays.

Case	Age and sex of the child	NSLT opinion
1	21 months – Girl	She walked at 19 months and she is very shy and inhibited.
		She was referred to motoric stimulation.
		Four month after the Gades evaluation, the observation process continues because she is still in process of adaptation.
2	18 months – Girl	She was brought to early attention. She was detected with a motor delay.
3	26 months – Girl	She had begun motoric treatment with 8 months.
		After Gades evaluation she started speech therapy treatment.
4	34 months – Girl	After Gades evaluation she started speech and language intervention.
5	39 months – Boy	After Gades evaluation he started speech and language intervention.
6	36 months – Boy	After Gades evaluation he started speech and language intervention.
7	42 months – Boy	After Gades evaluation he started speech and language intervention.
		The NSLT suggested that he is a child with family problems that may have affected the delay.

[a]NLST: nursery school language therapist.

Knowledge Base Accuracy

The NLST accepted Gades' decisions in 93% (87/94) of the 0-3 years cases and 67% (35/52) at the 4-6 years stage. Figure 3 shows that disagreements with Gades' decisions are higher in the 4-6 years stage where the NSLT indicated that some of the KB questions should be reviewed. The results comparison led to a total accuracy ratio for Gades KB of 84% (122/146). A total of 24 cases from the 146 sample set a 16% non-acceptance ratio to be reduced with further KB refinement. The NLST and experts consider that a golden pattern for Gades KB accuracy of 95% will be achieved after the ongoing review of pragmatic and language articulation items.

Figure 4 compares the acceptance of Gades' decisions in years 1 to 6. The area representing the therapist's agreement with the GADES system is greater than the area that expressed her disagreement. Years 1 and 3 reached a 100% acceptance ratio and year 5 up to 95%. The NSLT did not accept some Gades decisions at the fourth and sixth year, due to lack of agreement with some of the KB questions. The NSLT acceptance ratio of

80% of Gades decisions at year 2 has led to a recent update of the KB to enhance the system outcomes at this stage.

The accuracy of the KB questions after the language evaluation process at Legamar is grouped by year in Table 4. The second and third columns show the age range in months and the corresponding number of evaluated questions for each age range. The KB should have a minimum and necessary set of questions. The group of experts, who participated in the KB construction, stated that a range of 3-8 questions per month may be enough to achieve early detection of language disorders. Thus, the desired maximum number of questions in this range would be 48 (8 questions per month as a maximum) and this value is not reached in any group. The current version of the KB has a small number of questions for each month.

There are more questions at the 0-3 years stage, because early detection of language disorders is critical at this developmental period. The child evolves very quickly at this stage and the KB requires higher accuracy to analyze the evolution status. There are not questions for all months in the 4-6 years stage, because the therapists determined during the process of KB construction different age ranges to support a specific assessment. Questions are structured according to evaluative items at 42, 45, 46, 48, 54, 60, 66, 69, and 72 months.

The fourth column of Table 4 indicates the language development areas, evaluated by the KB questions. Finally, the last column details the opinion of the NSLT about the evaluated areas: correct questions or questions to be added, according to the KB for each year.

The NSLT indicates that the separation between speech perception and pragmatic is minimal. Besides, pragmatic disorders often coexist with other language problems such as vocabulary development or grammar. Pragmatic problems have lower social acceptance. The NSLT considers that the correct evaluation of pragmatics is important to avoid, or to treat as early as possible, a future neurological disorder.

Table 4. Accuracy and refinement of the knowledge base questions.

Year of the questions	Months of the questions	Number of KB questions	Evaluated areas	NSLT[a] opinion
Year 1	0-6	18	SR[b] - SP[c] - P[d]	Questions OK
	7-12	23	SP - SPD[e] - P	Questions OK
Year 2	13-18	17	SP - SPD - P	Questions OK
	19-24	18	SP - SPD - P	Disagree with some questions
Year 3	25-30	13	SP - SPD	Add questions of P
	31-36	10	SP - SPD	Add questions of P
Year 4	37-42	5	SP - SPD - P	In SP more questions of articulation language More questions of P
	43-48	11	SP - SPD - P	Questions OK
Year 5	49-54	3	SP - SPD	In SP more questions of articulation language Add questions of P
	55-60	4	SP - SPD - P	Add questions of P
Year 6	61-66	3	SP - SPD - P	Add questions of SP and SPD
	67-72	9	SP - SPD - P	Questions OK

[a]NSLT: nursery school language therapist

[b]SR: sensory reception

[c]SP: speech perception

[d]P: pragmatic

[e]SPD: speech production

Figure 3. Knowledge base accuracy by stage.

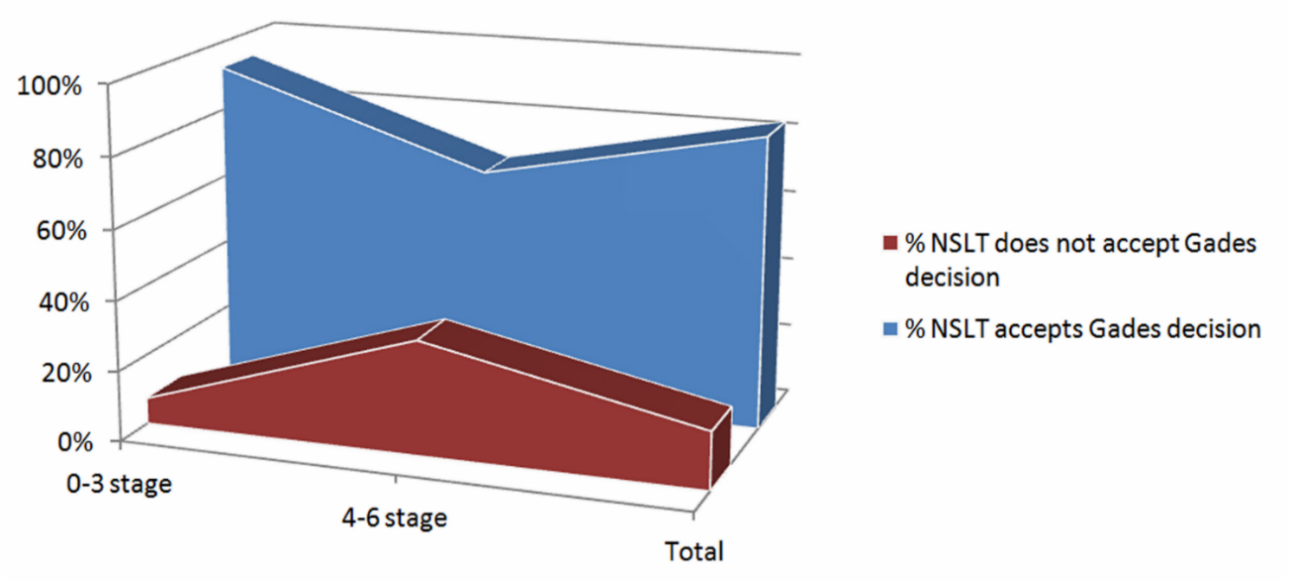
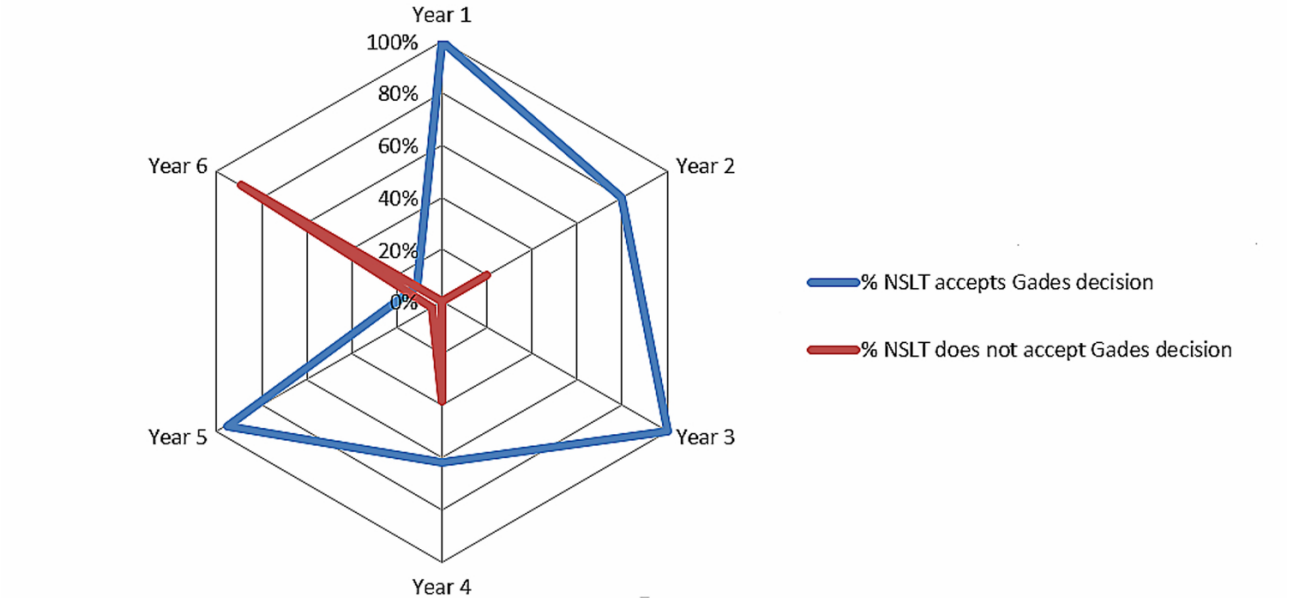


Figure 4. Knowledge base accuracy by age of the child.

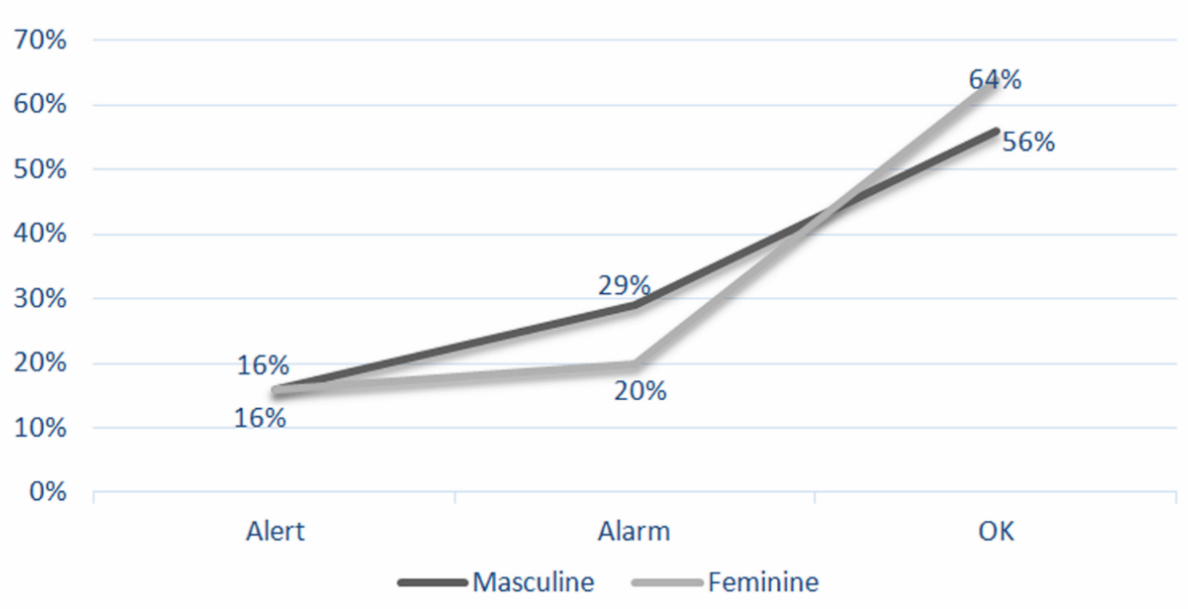


Evaluations Percentage by Result and Sex

Figure 5 illustrates that the percentage of alerts suggesting a pediatric visit is equal in both sexes. Despite having fewer language evaluations of boys than girls, it is remarkable to have a higher percentage of normative results (OK) for girls than for boys. Besides, most of the alarms, implying immediate referral to a specialist, occur for boys.

The language evaluation identified a total of 88 cases with a normative result (OK), a total of 35 cases with a referral to a specialist (Alarm), and a total of 23 with a follow-up pediatric visit (Alert). According to the sex of the child, a total of 83 girls were evaluated with the following results: 13 cases (16%) with Alarm, 17 cases (20%) with Alerts, and 53 (64%) cases with OK. A total of 63 boys were evaluated with the following results: 10 cases (16%) with Alarm, 18 cases (29%) with Alerts, and 35 cases (56%) with OK.

Figure 5. Evaluations percentage by result and sex.

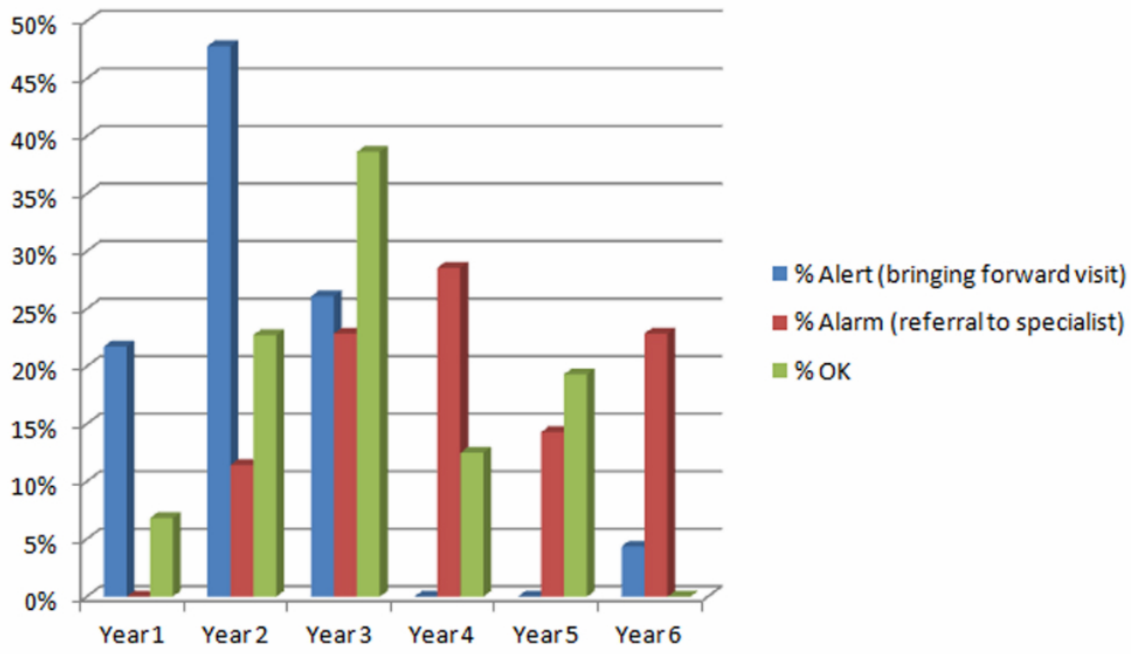


Evaluations Percentage by Result and Age

There were more language evaluation cases of children with normal development (OK) in the third year as Figure 6 illustrates. However, there were no abnormal language evaluations in the sixth year. A large percentage of language evaluations during the sixth year were referrals to a neuropediatrician or early attention (Alarm). This is due to the fact that there was a question that none of the children satisfactorily answered, which justifies the need to refine the KB before conducting new language evaluations with children or evaluations in the primary care real environment.

The maximum number of language evaluation cases of children with an Alert happened in the second year. The percentage of Alerts in years 1, 2, and 3 is higher than in years 4, 5, and 6. The higher percentage of Alarms at the 4-6 years stage is not significant because the NSLT detected some semantic mistakes in some KB questions. The current refinement of the KB is taking into account the opinion of the group of experts who originally participated in the KB construction and the evaluation results of Gades presented in this work. For this reason, the KB questions will not be reviewed until they can include the review from all experts, according to the analysis of the results of the Gades evaluation process summarized in Tables 2 and 3. After this, the modifications suggested for the KB refinement process will be adopted and a second evaluation process will be triggered with two more enrolled schools. Thus, some enhancements are expected in the NSLT acceptance of Gades according to the suggested refinement proposed to the system for years 4, 5, and 6.

Figure 6. Evaluations percentage by result and age.



Discussion

Principal Findings

Although previous studies relate positive experiences about the educator role in the process of language acquisition [4,31-38], the real impact of CDSS needs a deeper discussion. The adequate formalization of knowledge acquisition about associated symptoms in certain contexts conditions the reliability of Gades ontology, created for early detection of language disorders. The results obtained in this research, acquired from the input of 12 educators at the Legamar school, are consistent with the approach of Ygual-Fernandez et al, which supports feedback from nursery school educators to validate the decisions, as triggered by Gades. System outcomes validated in this research do align with recent studies like [39], which reinforce the use of computers, handheld, and mobile devices to provide instant access to extensive amounts and types of suitable information for health care professionals.

Language evaluation performed by Gades is consistent with the higher incidence of language impairment between boys and girls identified in scientific studies [3,40]. This research did not start from an equal ratio of male and female populations because it assumed the unbalanced gender distribution of the whole population of the targeted nursery school. However, the size of the sample (146 children) provided Gades with the capacity to support a significant aggregation of the results for both age and sex. The continuity of the training period, carried out for 3 months by the speech therapist with the educators, positively influenced their acceptance of Gades procedures and observations. This step-by-step methodological approach helped to foster the results obtained, which were ultimately able to successfully indicate the existence of language disorders in children.

The higher accuracy of Gades at 0-3 years than at the 4-6 years stage is directly related to the research rationale that gave priority to earlier and more precise detection of language disorders from 1 to 3 years old. This outcome does not invalidate the use of the system from 4 to 6 years old but paves the way for better performance at this stage after the KB refinement. Furthermore, the comparison between Gades' results and the expert's feedback shows that linguistic functionality from 30 months cannot be clearly formalized though a specific item. The NSLT pointed out the need to enhance the analysis of pragmatic skills at each contextual scenario from 3 years onward. The deeper specific review of Gades KB, including more specific questions of pragmatic evaluation, will require a second wide scale evaluation to measure the improvements of effectiveness and reliability for language evaluation results in the 4-6 years stage.

The definition of two incremental phases in the evaluation method of Gades' capacity for early detection of language disorders helped to provide users with requested information packages at the stage of need. The design of traditional questionnaires for parents and teachers, adapted from Gades KB, made it easier to assess the language use of the child in different interaction environments [32]. For example, the information offered by parents and teachers through the Children's Communication Checklist (CCC) demonstrated good sensibility and determined pragmatic difficulties that children might present such as autism, attention-deficit/hyperactivity disorder (ADHD), conduct disorder, Williams's syndrome, and Down's syndrome [4].

Limitations

The early detection of language disorder tools has well-documented limitations in the specialized literature [4], such as: (1) subjectivity of the person who completes the questionnaire or scale of values [33], and his/her previous

knowledge or specific training in relation to linguistic skills; (2) inconsistencies between the teacher's observations and the child's capacity in evaluation tests [34,35], due to possible differences in the child's linguistic conduct in spontaneous daily situations and to his/her execution during a formal evaluation, characterized by a major inflexibility [36]; and (3) trustworthiness of the predictive power of the questionnaires used due to the fact that they depend on the age of the children, where the estimations of teachers seem to be less trustworthy when smaller children are evaluated, in relation to the rapid cognitive and behavioral changes that they try out in these early ages [37,38].

The limitations of the aforementioned issues do not invalidate the current research as a high number of studies verified the existence of a significant correlation between the observations of teachers with diverse linguistic skills and the punctuations obtained by his/her pupils in different standardized evaluation tests. All of them used questionnaires focused on general or punctual aspects of linguistic processing as Gades inputs did [4]. Other works have also reported teacher difficulty in the detection of speech difficulties and a lack of sensibility for the differentiation of difficulties in the speech domain in every evolutionary moment [32].

Although the Legamar school is a private entity, its demographic data show a realistic potential to scale up the trials and results of this research to other public and private centers. It assists children from middle-class families with a normal distribution of gender, age, and parental income. The extrapolation of the study to other classes in public or private schools does not require methodological changes or a team of professionals to be involved. A higher number of children with language delays is expected to be obtained in a center where speech therapists do not belong to the regular staff. If the school does not have an NSLT, the method presented in this study cannot be applied equally and the NSLT functionality could fall to educators or another professionals.

Consulted language experts stated that the extrapolation of the study to another region where other dialects or languages are spoken may obtain similar results. Children all over the world learn more than one language without developing speech or language problems. Even though bilingual children develop language skills just as other children do [41], the introduction of a second language may slightly delay the acquisition timescale. This "silent period" can sometimes last several months. This is a normal evolution and the child will recover the proper developing stage [41].

Finally, we have not detected false negatives in any stage (0-3 or 4-6 years). The false positive rate in stage 0-3 years has been low. However, we had a high ratio of false positives in the 4-6 years stage. The NSLT detected the main causes of this ratio to be related to semantic mistakes in the questions involved. We are currently in a refinement process to solve this situation.

Conclusions

This research details an innovative solution to support knowledge-based detection of language disorders in children aged 0 to 6 years at nursery schools. The solution provides nursery school educators with a monitoring tool to assess the degree of language acquisition in their students. In spite of the additional workload faced by the educators, the school highlights the benefits of this type of monitoring for children.

The results of the evaluation at the Legamar Nursery School show that several children identified by Gades as having a possible language delay had not previously caused alarm to either the school therapist or to his/her educator. Further, a large number of children identified by Gades were also identified by the NSTL, especially in the 0-3 years stage. These results lead us to conclude that this kind of Web-based CDSS can undertake early detection of language delays in children at a nursery school with the support of their teachers, thus improving the neurodevelopmental follow-up.

In the process of early detection of language disorders, it is necessary to have not only a very specific knowledge, but also, a capacity for suitable observation. Therefore, we can summarize that Gades can be an effective CDSS for use by speech therapists and school psychologists in the rapid detection of children with difficulties in language development; Gades guides educators in the observation required for detection and also promotes the stimulation of skills aimed at diminishing and even preventing the appearance of these disorders; and Gades can be a collaboration tool involving parents and primary care pediatricians in the process of language evaluation.

Other conclusions of this research suggest the need to include supervised learning capacities in Gades. The learning functionality requires the definition of a specific model that allows a proper mix of automation and experts' supervision. Experts will be able to update Gades KB easily, taking into account the suggestions triggered automatically by the system. These suggestions will come from significant samples of real use cases. The following complementary proposed actions will improve the capacity of Gades detection in order to promote better health status of children. First, questions related to difficulties in the sound articulation domain can be incorporated. There can be situations in which children do not have a problem in language development, rather the problem derives from difficulties with sound discrimination. These questions will be studied by a multidisciplinary team of experts skilled in the relevant areas. Second, complementary evaluation in other areas outside of the school can be included. The observation capacity of the teachers, though it is considerable, does not include all aspects that would be desirable at the time of establishing a proper diagnosis. To be able to analyze other contexts outside of school, there is a version being adapted for primary care pediatricians and analysis has also begun for a possible version for parents. Third, new questions to improve aspects in the language domain can be refined and added. In the 4-6 years stage, the need for major refinement was detected in questions related to the pragmatics. Currently, we are working on it with the NSLT at Legamar.

Furthermore, the authors have defined a new concept called "Internet of Toys". It deals with the possibility of obtaining information about the child's development through his or her natural interaction with toys. This new interaction paradigm might provide Gades with the capacity to acquire real-time data

in order to improve its reasoning performance. Thus, the system could improve its effectiveness thanks to the very earliest utilization of information related to the behavior of the child. Data monitored via the expected interaction of children with certain toys could enhance Gades' reliability with more critical information.

Acknowledgments

The authors wish to thank María Peñafiel, principal at Legamar School (Madrid, Spain); Mª Teresa Ferrando, MD, neuropediatrician at the Quirón Hospital of Madrid; Paloma Tejeda of the Language Intervention Centre (LIC) Universidad Autónoma de Madrid; and José Arizcun, MD, expert neonatologist in developmental disorders and child disability.

This article is part of research we're conducting in the Talisec+ project (a framework for knowledge-based management of accessible security guarantees for personal autonomy; TIN2010-20510-C04-01), supported by the Ministry of Education and Science of Spain through the National Plan for R+D+I (research, development, and innovation).

Conflicts of Interest

None declared.

References

1. Grupo de atención temprana. Protocolo de observacion del lenguaje para maestros de educación infantil. Eficacia en la detección de dificultades semanticas y morfosintacticas. In: Real patronato de prevención y de atención a personas con minusvalía. Real Patronato sobre Discapacidad. Ministerio de Trabajo y Asuntos Sociales. Madrid, Spain (Spanish). Madrid: Centro español de documentación sobre discapacidad; 2005:405-420.
2. Council on Children With Disabilities, Section on Developmental Behavioral Pediatrics, Bright Futures Steering Committee and Medical Home Initiatives for Children With Special Needs Project Advisory Committee. Identifying infants and young children with developmental disorders in the medical home. Pediatrics 2006;118:405-420. [doi: 10.1542/peds.2006-1231]
3. Wilson P, McQuaige F, Thompson L, McConnachie A. Language delay is not predictable from available risk factors. ScientificWorldJournal 2013;2013:1-8 [FREE Full text] [doi: 10.1155/2013/947018] [Medline: 23576912]
4. Ygual-Fernández A, Cervera-Mérida JF, Baixauli-Fortea I. Protocolo de observacion del lenguaje para maestros de educacion infantil. Eficacia en la deteccion de dificultades semanticas y morfosintacticas. Rev Neurol 2011;52:127-128.
5. Prior M, Bavin E, Ong B. Predictors of school readiness in five - to six - year - old children from an Australian longitudinal community sample. Educational Psychology 2011 Jan;31(1):3-16. [doi: 10.1080/01443410.2010.541048]
6. Ripley K, Yuill N. Patterns of language impairment and behaviour in boys excluded from school. Br J Educ Psychol 2005 Mar;75(Pt 1):37-50. [doi: 10.1348/000709905X27696] [Medline: 15831180]
7. Lewis BA, Freebairn LA, Taylor HG. Academic outcomes in children with histories of speech sound disorders. J Commun Disord 2000;33(1):11-30. [Medline: 10665511]
8. Miniscalco C, Nygren G, Hagberg B, Kadesjö B, Gillberg C. Neuropsychiatric and neurodevelopmental outcome of children at age 6 and 7 years who screened positive for language problems at 30 months. Dev Med Child Neurol 2006 May;48(5):361-366. [doi: 10.1017/S0012162206000788] [Medline: 16608544]
9. Law J, Rush R, Schoon I, Parsons S. Modeling developmental language difficulties from school entry into adulthood: literacy, mental health, and employment outcomes. J Speech Lang Hear Res 2009 Dec;52(6):1401-1416. [doi: 10.1044/1092-4388(2009/08-0142)] [Medline: 19951922]
10. Paul R, Norbury CF. Language Disorders from Infancy through Adolescence: Listening, Speaking, Reading, Writing, and Communicating, 4th ed;. The Netherlands: Elsevier Science Health Science Division; 2011.
11. Paul R. Language Disorders from Infancy through Adolescence: Assessment and Intervention, 3rd ed;. Philadelphia: Elsevier Health Sciences; 2007.
12. Castro-Rebolledo R, Giraldo-Prieto M, Hincapié-Henao L. Trastorno específico del desarrollo del lenguaje: Una aproximación teórica a su diagnóstico, etiología y manifestaciones clínicas. Rev Neurol 2004;39:1173-1179.
13. Martín-Ruiz ML, Valero Duboy MA, Pau de la Cruz I. Deployment and validation of a smart system for screening of language disorders in primary care. Sensors 2013;13:7522-7523. [doi: 10.3390/s130607522]
14. World report on disability. Geneva. World Health Organization, World Bank URL: http://www.who.int/disabilities/world_report/2011/en/ [accessed 2014-03-20] [WebCite Cache ID 6ODG9eOmq]
15. Narbona J, Chevrie-Muller C. El lenguaje del nino. Desarrollo normal, evaluación y trastornos. Barcelona (Spain): Masson; 2003.
16. Law J, Boyle J, Harris F, Harkness A, Nye C. Screening for speech and language delay: a systematic review of the literature. Health Technol Assess 1998;2(9):1-184 [FREE Full text] [Medline: 9728296]
17. Law J. The Encyclopedia of Language and Literacy Development. London, Ontario: The University of Western Ontario; 2009. Short- and long-term outcomes for children with Primary Language Impairment (PLI) URL: http://literacyencyclopedia.ca/index.php?fa=items.show&topicId=263

18. Tomblin JB, Records NL, Buckwalter P, Zhang X, Smith E, O'Brien M. Prevalence of specific language impairment in kindergarten children. J Speech Lang Hear Res 1997 Dec;40(6):1245-1260. [Medline: 9430746]
19. Gillberg C. The ESSENCE in child psychiatry: Early Symptomatic Syndromes Eliciting Neurodevelopmental Clinical Examinations. Res Dev Disabil 2010;31(6):1543-1551. [doi: 10.1016/j.ridd.2010.06.002] [Medline: 20634041]
20. Ringwalt S. FPG Child Development Institute, National Early Childhood Technical Assistance Center. Chapel Hill: The University of North Carolina; 2008. Developmental screening and assessment instruments with an emphasis on social and emotional development for young children ages birth through five URL: http://www.nectac.org/~pdfs/pubs/screening.pdf [accessed 2014-05-26] [WebCite Cache ID 6PrZRe4W9]
21. The Early Intervention Program: A Parent's Guide. URL: http://www.health.ny.gov/publications/0532/ [accessed 2014-03-20] [WebCite Cache ID 6ODFumlAt]
22. Virginia Department of Education Division of Special Education and Student Services. Speech-Language Pathology Services in Schools: Guidelines for Best Practice. Richmond, VA: Virginia Department of Education; 2011:1-132.
23. Illingworth RS. The Normal Child: Some Problems of the Early Years and Their Treatment. In: Churchill Livingstone. UK: Churchill Livingstone; 1983.
24. Peralta MF, Narbona J. Retrasos del desarrollo verbal y problemas de aprendizaje escolar: estudio longitudinal. Bordón. Revista de Pedagogía 1991;43:285-298.
25. Botting N, Conti-Ramsden G, Crutchley A. Concordance between teacher/therapist opinion and formal language assessment scores in children with language impairment. Eur J Disord Commun 1997;32(3):317-327. [Medline: 9474284]
26. Miller RA. Medical diagnostic decision support systems--past, present, and future: a threaded bibliography and brief commentary. J Am Med Inform Assoc 1994;1(1):8-27 [FREE Full text] [Medline: 7719792]
27. Berner ES, Maisiak RS. Influence of case and physician characteristics on perceptions of decision support systems. J Am Med Inform Assoc 1999;6(5):428-434 [FREE Full text] [Medline: 10495102]
28. Martin-Ruiz ML, Duboy Valero MA, Pau de la Cruz I. Development of a Knowledge Base for smart screening of language disorders in primary care. In: 2012 IEEE 12th International Conference on Bioinformatics & Bioengineering (BIBE). 2012 Presented at: BIBE; Nov. 11-13, 2012; Larnaca p. 121-126. [doi: 10.1109/BIBE.2012.6399660]
29. Alonso Betanzos A, Guijarro Berdiñas B, Lozano Tello A, Palma Méndez JT, Taboada Iglesias M. Ingeniería del conocimiento. Aspectos Metodológicos. Madrid (Spain): Pearson Educacion; 2004:1-488.
30. Frankenburg WK, Dodds JB, Archer P. The DENVER II: A major revision and standardization of the denver developmental screening test. Pediatrics 1992;89:91-97.
31. Jessup B, Ward E, Cahill L, Keating D. Teacher identification of speech and language impairment in kindergarten students using the Kindergarten Development Check. Int J Speech Lang Pathol 2008;10(6):449-459. [doi: 10.1080/17549500802056151] [Medline: 20840024]
32. Bishop DV, Baird G. Parent and teacher report of pragmatic aspects of communication: use of the Children's Communication Checklist in a clinical setting. Dev Med Child Neurol 2001 Dec;43(12):809-818. [Medline: 11769267]
33. Bleses D, Vach W, Jørgensen RN, Worm T. The internal validity and acceptability of the Danish SI-3: a language-screening instrument for 3-year-olds. J Speech Lang Hear Res 2010 Apr;53(2):490-507. [doi: 10.1044/1092-4388(2009/08-0132)] [Medline: 20360468]
34. Antoniazzi D, Snow P, Dickson-Swift V. Teacher identification of children at risk for language impairment in the first year of school. Int J Speech Lang Pathol 2010 Jun;12(3):244-252. [doi: 10.3109/17549500903104447] [Medline: 20433343]
35. Geurts H, Embrechts M. Pragmatics in pre-schoolers with language impairments. Int J Lang Commun Disord 2010;45(4):436-447. [doi: 10.3109/13682820903165685] [Medline: 19821791]
36. Gilmore J, Vance M. Teacher ratings of children's listening difficulties. Child Language Teaching and Therapy 2007;23:133-156. [doi: 10.1177/0265659007073876]
37. Kenny DT, Chekaluk E. Early reading performance: a comparison of teacher-based and test-based assessments. J Learn Disabil 1993 Apr;26(4):227-236. [doi: 10.1177/002221949302600403] [Medline: 8515187]
38. Williams C. Teacher judgments of the language skills of children in the early years of schooling. Child Language Teaching and Therapy 2006;22:135-154. [doi: 10.1191/0265659006ct304oa]
39. Mickan S, Tilson JK, Atherton H, Roberts NW, Heneghan C. Evidence of effectiveness of health care professionals using handheld computers: a scoping review of systematic reviews. J Med Internet Res 2013;15(10):e212 [FREE Full text] [doi: 10.2196/jmir.2530] [Medline: 24165786]
40. Choudhury N, Benasich AA. A family aggregation study: the influence of family history and other risk factors on language development. J Speech Lang Hear Res 2003 Apr;46(2):261-272 [FREE Full text] [Medline: 14700370]
41. Learning two languages.: American Speech-Language-Hearing Association URL: http://www.asha.org/public/speech/development/bilingualchildren/ [accessed 2014-03-27] [WebCite Cache ID 6OO82kMk1]

Abbreviations

ADHD: Attention-Deficit/Hyperactivity Disorder
CCC: Children's Communication Checklist

CDSS: clinical decision support system
CK: CommonKADS
KB: knowledge base
LIC: Language Intervention Centre
NSLT: nursery school language therapist
P: pragmatic
QoLrH: Quality of Life related to Health (QoLrH)
SP: speech perception
SPD: speech production
SR: sensory reception
WHO: World Health Organization

Edited by G Eysenbach; submitted 21.01.14; peer-reviewed by C Bescos, V Traver; comments to author 13.03.14; revised version received 11.04.14; accepted 19.05.14; published 28.05.14.

Please cite as:
Martín Ruiz ML, Valero Duboy MÁ, Torcal Loriente C, Pau de la Cruz I
Evaluating a Web-Based Clinical Decision Support System for Language Disorders Screening in a Nursery School
J Med Internet Res 2014;16(5):e139
URL: http://www.jmir.org/2014/5/e139/
doi:10.2196/jmir.3263
PMID:24870413

Original Paper

Experiences of Multidisciplinary Development Team Members During User-Centered Design of Telecare Products and Services: A Qualitative Study

Joan Vermeulen[1*], MSc; Renée Verwey[1,2*], RN, MSc; Laura MJ Hochstenbach[1*], MSc; Sanne van der Weegen[1*], MSc; Yan Ping Man[2*], RN, MSc; Luc P de Witte[1,2*], MD, PhD

[1]CAPHRI, Department of Health Services Research, Maastricht University, Maastricht, Netherlands

[2]Research Center for Technology in Care, Faculty of Health Care, Zuyd University of Applied Sciences, Heerlen, Netherlands

[*]all authors contributed equally

Corresponding Author:
Joan Vermeulen, MSc
CAPHRI
Department of Health Services Research
Maastricht University
PO Box 616
Maastricht, 6200 MD
Netherlands
Phone: 31 433882199
Fax: 31 433884162
Email: j.vermeulen@maastrichtuniversity.nl

Abstract

Background: User-centered design (UCD) methodologies can help take the needs and requirements of potential end-users into account during the development of innovative telecare products and services. Understanding how members of multidisciplinary development teams experience the UCD process might help to gain insight into factors that members with different backgrounds consider critical during the development of telecare products and services.

Objective: The primary objective of this study was to explore how members of multidisciplinary development teams experienced the UCD process of telecare products and services. The secondary objective was to identify differences and similarities in the barriers and facilitators they experienced.

Methods: Twenty-five members of multidisciplinary development teams of four Research and Development (R&D) projects participated in this study. The R&D projects aimed to develop telecare products and services that can support self-management in elderly people or patients with chronic conditions. Seven participants were representatives of end-users (elderly persons or patients with chronic conditions), three were professional end-users (geriatrician and nurses), five were engineers, four were managers (of R&D companies or engineering teams), and six were researchers. All participants were interviewed by a researcher who was not part of their own development team. The following topics were discussed during the interviews: (1) aim of the project, (2) role of the participant, (3) experiences during the development process, (4) points of improvement, and (5) what the project meant to the participant.

Results: Experiences of participants related to the following themes: (1) creating a development team, (2) expectations regarding responsibilities and roles, (3) translating user requirements into technical requirements, (4) technical challenges, (5) evaluation of developed products and services, and (6) valorization. Multidisciplinary team members from different backgrounds often reported similar experienced barriers (eg, different members of the development team speak a "different language") and facilitators (eg, team members should voice expectations at the start of the project to prevent miscommunication at a later stage). However, some experienced barriers and facilitators were reported only by certain groups of participants. For example, only managers reported the experience that having different ideas about what a good business case is within one development team was a barrier, whereas only end-users emphasized the facilitating role of project management in end-user participation and the importance of continuous feedback from researchers on input of end-users.

Conclusions: Many similarities seem to exist between the experienced barriers and facilitators of members of multidisciplinary development teams during UCD of telecare products and services. However, differences in experiences between team members

from various backgrounds exist as well. Insights into these similarities and differences can improve understanding between team members from different backgrounds, which can optimize collaboration during the development of telecare products and services.

(J Med Internet Res 2014;16(5):e124) doi:10.2196/jmir.3195

KEYWORDS

user-centered design; telecare; eHealth; participation; multidisciplinary team; barriers and facilitators

Introduction

The number of people who suffer from chronic conditions and the number of elderly people are increasing, and at the same time, the number of care professionals is decreasing [1]. The gap between the demand for and supply of care that results from these changes will put a burden on patients, care professionals, and health care systems in the near future [2,3]. Telecare products and services have the potential to alleviate this burden by supporting self-management, remote monitoring of health conditions, and the delivery of interventions [4-6]. The appropriate adoption of telecare technologies can contribute to the lives of patients with chronic conditions and elderly people by improving their quality of life and enabling them to live independently for as long as possible [7,8].

Unfortunately, previous studies show that telecare does not always fulfill its potential because the telecare products and services that are developed do not fit the needs and preferences of end-users or because they do not fit the context in which they should be implemented [9-12]. Involving end-users in the development of telecare products and services can ensure that human and non-technology issues are taken into consideration, which improves the usability and acceptability of the technology that is being developed [13-15].

Theoretical frameworks regarding user-centered design (UCD) have been used as guidelines during the development processes of telecare products and services [16-22]. Such frameworks take into account important key principles of UCD, for example, the development process should be iterative and incremental, the end-users should be actively involved from early development stages onwards, design options should be explained to end-users in a language that they understand, the developed services and products should be evaluated in a real life context, and the development process should be performed by effective multidisciplinary teams [23]. Previous studies revealed that various barriers and challenges can occur during UCD processes that can influence the collaboration between multidisciplinary stakeholders. Examples of such barriers are communication between team members from different backgrounds, management of expectations, and availability of recourse for the involvement of users [14,24,25].

Understanding how members of multidisciplinary development teams with different backgrounds experience the UCD process might help gain insight into factors that different members consider critical during the development of telecare products and services. Furthermore, it might reveal how they deal with barriers and challenges they encounter. Therefore, the primary objective of this study was to explore how members of multidisciplinary development teams experienced the UCD process of telecare products and services. The secondary objective was to identify differences and similarities in the barriers and facilitators they experienced.

Methods

Design and Participants

This qualitative study was conducted in parallel with four (subsidized) Research and Development (R&D) projects that were initiated by Maastricht University and/or Zuyd University of Applied Sciences. The projects were selected because they developed a variety of telecare products and services to support self-management in different user-groups: patients with diabetes or chronic obstructive pulmonary disease (COPD) [26-28], patients with cancer pain, frail elderly people, or elderly people living at home [29]. A prerequisite for project selection was that end-users were considered official members of the development team. Multimedia Appendix 1 provides information regarding the four projects. Principle investigators of the projects had a background in health sciences.

For this study, a purposeful sample of the “core members” of the four development teams, who were involved throughout the whole development process from the beginning onwards, were invited to participate in a semistructured interview. Principle investigators of each development team identified the core members in consultation with their team and asked them whether they would like to receive an invitation for this study via email. The 25 core members were identified across the four R&D projects: 3 elderly persons, 2 professional advisors of these elderly persons (specialized in facilitation participation of elderly persons during research projects), 2 patients with chronic conditions, 1 geriatrician, 2 nurses, 5 software/technical engineers, 4 managers of R&D companies or engineering teams, and 6 researchers/principle investigators of the R&D projects. All these potential participants received an invitation letter via email from the researcher (JV) explaining the purpose and details of this qualitative study. All invitees accepted the invitation.

Procedures

After invitees agreed to participate, the interview was scheduled at a place and time that was convenient for the participant. As a result, 13 interviews were conducted at the university, nine at the workplace of the participant, and 3 via Skype/telephone call. The 6 participants of this study who were researchers or principle investigators in one of the four R&D projects (JV, RV, LMJH, SvdW, YPM, LdW) were interviewed by an experienced external interviewer who was not involved in any of the R&D projects described above. JV, RV, LMJH, SvdW, and YPM interviewed the remaining 19 participants of this study. All participants were interviewed by a researcher who

was not involved in the same R&D project as the participant. Interviews with members of development teams of Projects 1, 3, and 4 were conducted when final products were already developed. Interviews with members of the development team of Project 2 were conducted at a prototype stage. Interviews were recorded using a digital voice recorder and lasted between 35 and 65 minutes. Average duration of an interview was 46 minutes.

All semistructured interviews were based on the following topic list: the role of the participant in the project, the aim of the project, experiences during the development process, points of improvement for the development process, and what the project meant to the participant. This topic list was established before the start of the study and was not changed during the course of the study. Interviewers used reflective listening techniques to encourage participants to elaborate on their views. By reflecting back to the participants during the interview what the interviewer believed was said, the statements of the participant were verified and/or clarified. By keeping an open posture and by emphasizing at the beginning of the interview that there were no right or wrong answers, participants were encouraged to share their own personal opinions and thoughts. All participant answers were kept confidential. To guarantee anonymity, the following terms are used in this article to refer to the data obtained from certain groups of participants:

- "End-user" refers to data obtained from an elderly person, an advisor of an elderly person, or a patient with a chronic condition (n=7).
- "Professional end-user" refers to data obtained from a geriatrician or a nurse (n=3).
- "Engineer" refers to data obtained from a software/technical engineer (n=5).
- "Manager" refers to data obtained from a technical project leader or owner of an R&D company (n=4).
- "Researcher" refers to data obtained from a researcher or principle investigator of the R&D projects (n=6).

Data Analyses

Once all interviews were conducted, they were transcribed verbatim by two of the researchers (JV/RV) or a research assistant. Afterwards, JV checked the transcripts against the audio recordings. All transcripts were coded using NVivo version 9.0. Field notes from the interviews were also included in the analyses if they were available. Two researchers (JV & RV) started analyzing the data using a conventional content analysis approach. They independently coded six transcripts of interviews that were conducted with members from different development teams with different backgrounds. After initial coding, the 2 researchers checked for consensus, and after discussion, they agreed on the main themes and subthemes of the coding scheme. This coding scheme was used by JV to analyze the remaining interviews. If in doubt about whether data from these remaining interviews fitted the coding scheme or not, JV consulted RV. Themes and subthemes were refined or extended based on the data from the remaining interviews to be analyzed, and if necessary, new (sub)themes were added. Once all transcripts were analyzed, the content of the themes and subthemes of the coding scheme were discussed with the research group that included all co-authors of this paper. Consensus was reached on themes that were related to different phases of the development process and subthemes that related to barriers and facilitators participants experienced during these phases. After completing the data analyses, findings related to themes and subthemes were reported back to all participants for a member check.

Results

Themes

The main themes that emerged from the analyses related to different phases of the UCD process were (1) creating a development team, (2) expectations regarding responsibilities and roles, (3) translating user requirements into technical requirements, (4) technical challenges, (5) evaluation, and (6) valorization. Experiences of participants during these different phases are described below. Tables 1 and 2 provide an overview of which experienced barriers and facilitators respectively were reported by which members of the development team, to gain insight into differences and similarities.

Table 1. Experienced barriers in the UCD process of telecare products and services according to different members of the development team.

		Participant groups[a]				
Theme	Barriers	EU	PU	EN	MA	RE
Creating a development team	Team members come from different backgrounds and therefore do not speak the same “language” (use different terminology).	X	X	X	X	X
Expectations regarding responsibilities and roles	Team members have different implicit expectations regarding project management, tasks of team members, and delivery of content (especially at the start of the project).	X	X	X	X	X
Translation of user requirements into technical requirements	Prioritizing user-requirements with various stakeholders is more time consuming than expected.			X	X	X
	Iterative adaptations of user-requirements (especially in later stages) place a serious strain on the budget/time of the project.			X	X	X
Technical challenges	Integration of different technologies or platforms into one telecare service is difficult (but necessary).	X	X	X	X	X
	Time allowed for telecare development is short in subsidized projects which causes problems with robustness in the real life setting and large scale evaluation research.					X
	The commercial market is developing similar products at a rapid pace which makes it difficult to keep up.	X		X	X	X
Evaluation	Members of the development team are not the best evaluators because they find their own “work-arounds” to avoid bugs (unaware).			X		X
	Recruitment of patients and professionals for the longitudinal evaluation of the developed telecare products/services is time-consuming.	X	X	X	X	X
	Too many different projects/devices are offered to potential end-users at the same time.				X	X
Valorization	Different partners/companies who are involved have different ideas about what makes a good business case.				X	

[a]EU=end-users, PU=professional end-users, EN=engineers, MA=managers, RE=researchers.

Table 2. Experienced facilitators in the UCD process of telecare products and services according to different members of the development team.

		Participant groups[a]				
Theme	Facilitators	EU	PU	EN	MA	RE
Creating a development team	Team members recognize their complementary knowledge and skills, which creates a team spirit.	X	X	X	X	X
	Team members agree that end-users should be involved in the development team from the beginning of the development process.	X	X	X	X	X
	Project leaders and managers advocate that input of all team members should be treated as equal.	X				
	Researchers report back to end-user representatives about why their advice was followed or not.	X				
	Researchers visit team members at home/work (when team meetings are not possible due to differences in schedules).	X	X	X		
Expectations regarding responsibilities and roles	Team members voice expectations at the start of the project to prevent miscommunication at a later stage (but this may be difficult due to the fact that many expectations are implicit).	X	X	X	X	X
Translation of user requirements into technical requirements	Engineers help researchers to translate user-requirements into technical requirements to speed up this process.			X	X	X
Technical challenges	Researchers should take enough time to conduct small scale usability tests and pilot studies before moving to large trials to improve technical functioning and robustness.					X
	Products/services developed in the projects are easier to integrate in care processes compared to off-the-shelf products.	X	X	X	X	X
Evaluation	Members of the development team evaluate prototypes in lab to identify bugs and/or gain insight into experiences with the products/services.	X		X		X
	Care professionals receive reimbursement for the increased workload that comes with participating in an evaluation study.		X			X
	Study participants can be recruited via the network of the members of the development team, especially via patient/elderly representatives.	X				X
Valorization	Allocate part of the budget to the development of a business case and start with this at the beginning of the project.			X	X	

[a]EU=end-users, PU=professional end-users, EN=engineers, MA=managers, RE=researchers.

Creating a Development Team

Due to the fact that members of the development teams had different backgrounds, there was also a difference in knowledge about the project, UCD processes, telecare products and services, intended end-users of the technology, and ways of conducting research in a health care setting. In addition, different members of the development team seemed to speak a "different language" (use different terminology):

> *We have researchers, we have doctors, we have technicians, and it is very, it is not the same world. Because we do not speak the same language…Sometimes we think something and for example the researcher understands something else…We speak technical, she (the researcher) talks with elderly people, she talks with other worlds. That is why it is so difficult.* [P20, male, engineer]

Participants reported that overcoming these differences was sometimes difficult. A benefit of these differences, mentioned by most participants, is that members of the team could really complement each other, which resulted in a positive "team spirit":

> *They (technical engineers) did not know anything about medication and then you think; well I don't know anything about computers. So together that was fun, as if you speak a different language but still have to come up with a solution together.* [P23, female, professional end-user]

Multidisciplinary collaboration was positively influenced by the fact that patients and care professionals were involved and recognized as members of the development teams from the start. When managers or supervisors advocated that the input of all members should be treated as equal, this facilitated co-creation and UCD, according to end-users. Furthermore, end-users indicated that they appreciated it when the researcher reported back to them about which parts of their advice were followed and which were not. According to some, this feedback was more important than actually following their advice because it made them feel appreciated as a team member:

> *I did have the feeling that, the things that we put forward, that they (the researchers) did something with that. And sometimes they (the researchers) just said: "listen, we did not choose this or we did choose this". And I think that is important…Feedback is very*

> *important. For example when you put something forward and after that you don't hear anything. Then you don't know whether something was done with it at all or whether it was taken seriously. And then it is also difficult to be in that process together.* [P9, female, end-user]

Finally, the organizational aspect of working in multidisciplinary teams was a challenge in most projects. Since most members had different schedules, it was often difficult to organize meetings with the whole team. Submeetings were a satisfactory solution according to most. Elderly representatives, care professionals, and engineers indicated that they appreciated the flexible attitude of researchers who visited them at home or at work to offer additional information/explanation or an update.

Implicit Expectations Regarding Responsibilities and Roles

Participants indicated that during the project they discovered differences in expectations regarding responsibilities and roles between members of the development team. Examples of issues where expectations differed between members of the development team were who the project manager is, what the tasks are of different team members, and who delivers content of the services that are being developed. Participants seemed to agree that clearly expressing and communicating expectations at the start of the project would help to ensure that the entire development team was on the same page, which would probably optimize multidisciplinary collaboration. However, some participants experienced that voicing these expectations could be difficult since they are often implicit:

> *Look, in fact you start this project with an open mind and you create the expectations along the way. Because we work in a certain way: iterative design. We take it a step further every time. So, at the beginning of the project you actually do not know where you will end up. You work towards the end gradually. So, you actually do not have very detailed expectations before the start.* [P15, male, engineer]

> *I expected more vision, more strategy, a clearer picture about this (concept of the project) at the start of the project. Well that disappointed me. That fit with the patient group was created along the way. I would have liked to feed of the available knowledge of the university. But that was disappointing.* [P16, male, manager]

> *I think that we had a very large part in delivering content (for the telecare product and service)…Of course, you have to partly develop the intervention. But this detailed, no I didn't expect that I would do that. I expected another share of the company and thought that was disappointing in hindsight.* [P2, female, researcher]

Translating User Requirements Into Technical Requirements

Participants experienced that collecting user requirements and discussing with the development team which requirements deserved priority took longer than expected. One reason for this is that reaching a consensus with a team consisting of members with different backgrounds can be quite challenging.

Another time-consuming part of the process was translating user requirements into the very detailed technical requirements that engineers need to be able to develop the first prototypes. Help from the engineers in this translation appeared to be crucial since most researchers did not have a technical background and were not familiar with the technical language that is used to describe technical requirements:

> *In the concept phase, everything seems possible. But when you have to specify things until the final feedback message, it is very difficult. And it takes a lot of time to think these things through.* [P5, female, researcher]

> *The researchers are more of less forced into a role in which they have to think along in a technical manner. And they are not used to that, it is not their job. And that creates a certain type of tension. Because they are forced to think about (technical) things that they had not thought about before.* [P18, male, engineer]

Finally, the identified user requirements evolved during the project as a result of the iterative nature of the UCD processes. Members of the development team agreed that these iterations were necessary to ensure that the developed products and services meet the requirements of the end-users (in the best possible way). Especially engineers and managers pointed out that they tried to be as flexible as possible in incorporating the new and additional requirements in new prototypes because they recognized the importance of the iterations. However, at a certain point, this flexibility ends because deadlines have to be met and personnel will be deployed in other projects after these deadlines:

> *It is good to get feedback from your target group but it is important to stay in control…I had the feeling that there were too many changes in response to feedback of the target group…At a certain time we froze the specifications and started working with that.* [P18, male, engineer]

> *We had a lot of backwards and forwards and changing. What might be a relatively simple change for an end-user, for instance the change of a bar from one place to another, could take a significant amount of time or require a major change in the way a software program was running.* [P19, male, manager]

Technical Challenges

Overview

The technical challenges that the development teams were faced with related to the integration of different technologies or platforms, robustness of technology in a real-life setting, and rapid pace of developments on the commercial market.

Integration of Different Technologies or Platforms

The telecare products and services that were developed during the projects all required the integration of various technological

components (eg, integration of a sensor with an application on a mobile device and an online database). Technical and software engineers, and to a lesser extent also other participants, experienced that the integration of these components was often more difficult than expected for several reasons. First, engineers experienced that the components sometimes have their own "language" and specifications (eg, hardware vs software), which makes integration more complicated. Second, the input for the system as a whole, and input regarding the integration of the developed technology in care, was provided by different members of the development team. All team members recognized that integrating the input from these parties into the products and services can be a real challenge as it does not always match. Third, in some cases the different components are not finished at the same time. This can seriously delay the progress of projects.

> *Every company developed their part (of the technology)…We develop something nice but then it does not fit. And then the other company develops something nice and our system cannot handle it. And then the researchers provide a new part of the content and then we think: where can we put this? That was difficult to work with sometimes. We resolved it in the end but I think things could have run more efficient in some areas.* [P17, male, engineer]

Robustness of Telecare Products in Real-Life Setting

Researchers experienced that there was often not enough time and budget for the development of the telecare products because funding bodies assume that the technology already exists. Technology development in these projects appeared to be more challenging since the health care setting was taken into consideration, which is often not the case in existing off-the-shelf telecare products. As a result, the technology was sometimes not "mature" enough when projects moved to large trials or evaluation studies, which caused problems regarding robustness. All members of the development teams agreed that robustness is an absolute precondition for the uptake of telecare products and services in practice. Taking enough time for conducting in-lab usability tests and pilot studies before moving to real life settings might prevent problems concerning robustness in large trials according to researchers:

> *Funding agencies for research often assume that the technologies already exist. However, there is hardly any time for the development of the technology. But we know from experience that this is very difficult and time consuming. We should not jump to large evaluation studies too fast but first do pilots and usability evaluations before we start with the big works. That is something that is often underestimated.* [P6, male, researcher]

Developments on the Commercial Market

Many participants experienced that the commercial market sets the standard for products developed in the R&D projects; the user requirements for the telecare products and services are often influenced by what is already on the market. According to members of the development teams, it is nearly impossible to keep up with the rapid pace at which the commercial market is developing. The reason the development teams chose to create something new instead of buying off-the-shelf products is that newly developed products can be adapted to fit the health care context and end-user needs.

Evaluation

Overview

In all projects, the developed devices were tested in the lab and in real-life. Issues that relate to the evaluation of the devices in both contexts are described below.

Evaluation in Lab

The developed products were first tested extensively by the members of the development team. Researchers who tested the products indicated that this was not only important for identifying bugs and errors but that it also provided better insight into experiences expected with end-users of the product. However, researchers and engineers indicated that eventually they were less able to identify bugs, as they unconsciously created their own "work-arounds":

> *Once you actually use it (the technology) yourself, then you really bond with it so to speak. And then you become even more alert to improvements* [P2, female, researcher]

Evaluation in Real-Life

Participants experienced that the recruitment of patients and professionals who were willing to evaluate the developed products and services was a true bottleneck in the planning of the projects. A possible explanation for this problem, according to the participants, was that usually care professionals have to perform the activities for such R&D projects (eg, providing input for requirements or recruitment of patients) on top of their regular activities, which causes an increased workload. In some projects, the professionals did not get reimbursed for recruiting participants. Another problem described was that there were too many different (telecare) research projects in one area at the same time. Possible solutions for these problems, employed during the R&D projects, were reimbursing care professionals, recruiting participants via the (in)formal network of members of the development team, or organizing meetings with professionals who participate in evaluation studies to increase their awareness and create commitment to the project:

> *Having patience is very important in research, especially when it concerns including patients in your study. You will come across barriers, it is just like hurdling in a sense.* [P24, male, professional end-user]

> *Convincing people who have to do it alongside their job. They are doctors, and this comes on top of it…Yes, I think that people are bombarded with something new every time: another technique that would be nice. That makes it pretty difficult.* [P14, male, manager]

Valorization

Once the effects of the developed telecare products and services have been evaluated, the subsidized R&D projects will end.

Members of all development teams emphasized the importance of scaling up, and valorization of, the developed telecare products and services:

> *That is one of the problems; it will have to happen on a larger scale in the future. For that you do need a business model. Just doing projects for the sake of doing projects and doing something else when the funding ends does not make sense.* [P22, male, manager]

Managers from the R&D companies and some of the technical engineers indicated that it was too bad that there was little time or budget to focus on the development of a proper business model during the projects. The fact that different companies and organizations are often involved in the delivery of one telecare product and/or service is an important issue in the development of such business models and therefore a concern to most of the managers:

> *At the basis, you have to start thinking who will be the owner and who will make money on this…Somehow you have to figure out a model that covers the costs and leaves something extra because we are not the type of entrepreneur that keeps investing in something that does not make any money. So I do see possibilities, but it has not been defined clearly yet.* [P14, male, manager]

Elderly people, patients with a chronic disease, care professionals, and researchers also indicated that they had concerns regarding the upscaling of the developed products and services. Their concerns focused on questions such as who would pay and sell the products/services, how much they should cost, and how roles and responsibilities of patients and professionals would change when telecare products and services were implemented in practice:

> *I am afraid that it will not be financially viable. Because physiotherapy is not included in the basic health insurance which makes your target group very small.* [P7, male, end-user]

> *I think that we still have to think about that, make agreements about when you reply or do not reply (to the patient after a certain signal from the telecare product), better agreements on how you communicate with the patients (using the new telecare services).* [P25, male, professional end-user]

Discussion

Principal Results and Comparison to Previous Research

The current study provides insight into how members of multidisciplinary development teams experience the UCD process of telecare products and services. Several barriers and facilitating factors were experienced that can influence the UCD process according to different members of four multidisciplinary development teams. Most barriers that were identified were in line with previous research [14,24,25]. Some of these factors were reported by participants from different backgrounds, whereas others seemed to be more specific for a group of participants who share a similar background or role in the UCD process. Furthermore, this study also provided insight into how members of development teams tried to deal with barriers that they encountered and which actions they undertook to facilitate the development process.

All participants experienced that differences in background can cause a language gap between members of the multidisciplinary development team that can negatively affect the development process. Other researchers and designers who have recognized this barrier recommend the use of personas, scenarios, mock-ups, or prototypes to reduce the language gap [13,21,30,31]. These techniques were also used during the R&D projects that were the focus of this study. In addition, this study revealed that the gap between members of a development team can be bridged by emphasizing the way that team members can complement each other through their differences, by principal investigators who explicitly advocate equality of team member's input, and by providing feedback to (professional) end-users regarding their input from the start onwards, throughout the development process.

Participants with different backgrounds all recognized that managing expectations regarding responsibilities and roles is a critical factor during the UCD of telecare products and services. Previous research by Gasson suggests that managing expectations regarding work roles and tasks is a critical issue in any UCD process [32]. The results of the present study revealed that although all members of the development team had read and agreed on the same project plan and UCD methodology to be used, differences in expectations still existed. Goodman-Dean et al explain that despite agreeing on the same basic nature of the UCD process, differences can still exist between the approaches of the team members [33]. Explicitly voicing detailed expectations at the start of the development process might prevent delay later in the project. However, participants of the current study recognized that this might be difficult to do. Previous studies have emphasized that project management should identify and allocate responsibilities, tasks, and roles [16,32,34]. Project managers should facilitate the UCD process in this way without being overly prescriptive or bureaucratic since that might impede the creative nature of the design process [34].

The results of this study confirmed that time- and budget-related issues seemed to play an important role during different phases of the UCD process. The main time constraining factor reported in previous studies is that working with users in an iterative way takes too much time, regardless of the methodology used [13,14,34,35]. A literature review by Shah et al suggests that the lack of available end-users is a barrier during the UCD process [14]. The findings of the current study are not entirely in line with this since recruiting users to be involved in the development team and to test prototypes during the iterative UCD process seemed a lesser barrier compared to recruiting participant for longitudinal research. A possible explanation for this could be that for the prior UCD activities, participants were recruited via the (informal) network of end-users and professional end-users who were part of the development team. Through this, the involvement of (professional) end-users from early stages onwards facilitated the progress of the projects.

Other studies that have been conducted previously did not indicate which factors were considered too time-consuming. The current study provides insight into causes of time-related barriers as experienced by different members of multidisciplinary development teams. Researchers and engineers were the groups of participants who most frequently reported on time-related barriers and how they tried to deal with them.

A critical factor reported on by all managers is the development of a good business model. Results of this study revealed that budget should be allocated to the development of a business case and that the stakeholders involved should discuss business modeling issues at an early stage. Previous studies have emphasized that business model development should run in parallel with the UCD of eHealth technologies because it contributes to the development of such technologies [16,36].

Strengths and Limitations

The scope of the four R&D projects included in this study varied from general services related to care and well-being for elderly people living in the community to specific services for severely ill patients who suffer from cancer (treatment) pain. The variety of projects included increases the generalizability of the findings. Furthermore, all core members of the multidisciplinary development teams were interviewed in order to incorporate all points of view in our analyses. This is a strength of the study design since previous studies aimed at identifying barriers of UCD processes often focused merely on the perspective of the designers or developers [13,33,36]. Including participants from different backgrounds created triangulation of data resources, which increased trustworthiness of the findings. Furthermore, the member check revealed that participants agreed with the experienced barriers and facilitators that were identified.

When interpreting the results of this study, some limitations of the study design should be taken into account. First, in the interviews, participants reflected on development processes that started 1-2 years ago, which might have caused some degree of memory bias. However, reflecting on the development process at later stages might also have benefits over interviewing participants in the middle of the development process. In the latter situation, the answers of participants may be influenced by the stage of the project at the time of the interview. Second, participants were asked to reflect critically upon development processes in which they were involved, and consequently they had to reflect on their own actions and the actions of their team members, which could have been a sensitive topic. In order to limit this sensitivity, a researcher who was not part of the participant's development team conducted the interview. However, it cannot be ensured that all barriers and facilitators were reported to the interviewers. Third, the authors of this paper had a double role in this study since they interviewed the participants who were not researchers, but they were also core members of the development teams (and hence were interviewed themselves). The involvement and experiences of the researchers with UCD of telecare products and services could have influenced data collection and/or analyses. We aimed to minimize these influences by letting researchers interview only participants who were involved in different R&D projects than they were, by developing a topic list for the interviews that was used by all interviewers throughout the study, and by developing the coding scheme with 2 researchers. Finally, data from the interviews with the researchers and from the interviews with other team members were treated as equal in the analyses. This is not necessarily a limitation, but it might be considered a notable feature of the study design. The main reason for this novel and somewhat unusual approach was that the researchers themselves had experienced and influenced the development process just as much as other members of the development team. Not interviewing the researchers could have resulted in missing barriers and facilitators that they themselves experienced.

Conclusions

Many similarities seem to exist between the barriers and facilitators experienced by members of multidisciplinary development teams during UCD of telecare products and services. However, differences in experiences between team members with different backgrounds exist as well. Insights into these similarities and differences can improve understanding between team members from different backgrounds, which can optimize collaboration during the development of telecare products and services.

Acknowledgments

The authors would like to thank all members of the development teams for agreeing to participate in this study. Furthermore, Inge van der Putten is acknowledged for her help in conducting and transcribing the interviews, and April Boessen is acknowledged for her help in transcribing the interviews.

This study was embedded in four R&D projects that were funded by the Netherlands Organization for Health Research and Development (ZonMw) or the Dutch Cancer Society (KWF).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Description of four R&D projects developing telecare products and services.

[PDF File (Adobe PDF File), 148KB - jmir_v16i5e124_app1.pdf]

References

1. United Nations, Department of Economic and Social Affairs, Population Division. World Population Ageing 2009. New York: United Nations; 2009. URL: http://www.un.org/esa/population/publications/WPA2009/WPA2009_WorkingPaper.pdf [accessed 2014-05-07] [WebCite Cache ID 6POKO6xnX]
2. Wiener JM, Tilly J. Population ageing in the United States of America: implications for public programmes. Int J Epidemiol 2002 Aug;31(4):776-781 [FREE Full text] [Medline: 12177018]
3. Howse K. Healthy ageing: the role of health care services. Perspect Public Health 2012 Jul;132(4):171-177. [doi: 10.1177/1757913912444805] [Medline: 22729007]
4. McLean S, Sheikh A, Cresswell K, Nurmatov U, Mukherjee M, Hemmi A, et al. The impact of telehealthcare on the quality and safety of care: a systematic overview. PLoS One 2013;8(8):e71238 [FREE Full text] [doi: 10.1371/journal.pone.0071238] [Medline: 23977001]
5. Eland-de Kok P, van Os-Medendorp H, Vergouwe-Meijer A, Bruijnzeel-Koomen C, Ros W. A systematic review of the effects of e-health on chronically ill patients. J Clin Nurs 2011 Nov;20(21-22):2997-3010. [doi: 10.1111/j.1365-2702.2011.03743.x] [Medline: 21707807]
6. Kreps GL, Neuhauser L. New directions in eHealth communication: opportunities and challenges. Patient Educ Couns 2010 Mar;78(3):329-336. [doi: 10.1016/j.pec.2010.01.013] [Medline: 20202779]
7. Botsis T, Hartvigsen G. Current status and future perspectives in telecare for elderly people suffering from chronic diseases. J Telemed Telecare 2008;14(4):195-203. [doi: 10.1258/jtt.2008.070905] [Medline: 18534954]
8. Koch S, Hägglund M. Health informatics and the delivery of care to older people. Maturitas 2009 Jul 20;63(3):195-199. [doi: 10.1016/j.maturitas.2009.03.023] [Medline: 19487092]
9. Broens TH, Huis in't Veld RM, Vollenbroek-Hutten MM, Hermens HJ, van Halteren AT, Nieuwenhuis LJ. Determinants of successful telemedicine implementations: a literature study. J Telemed Telecare 2007;13(6):303-309. [doi: 10.1258/135763307781644951] [Medline: 17785027]
10. Pagliari C. Design and evaluation in eHealth: challenges and implications for an interdisciplinary field. J Med Internet Res 2007;9(2):e15 [FREE Full text] [doi: 10.2196/jmir.9.2.e15] [Medline: 17537718]
11. Teixeira L, Ferreira C, Santos BS. User-centered requirements engineering in health information systems: a study in the hemophilia field. Comput Methods Programs Biomed 2012 Jun;106(3):160-174. [doi: 10.1016/j.cmpb.2010.10.007] [Medline: 21075471]
12. Heeks R. Health information systems: failure, success and improvisation. Int J Med Inform 2006 Feb;75(2):125-137. [doi: 10.1016/j.ijmedinf.2005.07.024] [Medline: 16112893]
13. Kujala S. User involvement: a review of the benefits and challenges. Behaviour & Information Technology 2003;22(1):1-16 [FREE Full text]
14. Shah SG, Robinson I. Benefits of and barriers to involving users in medical device technology development and evaluation. Int J Technol Assess Health Care 2007;23(1):131-137. [doi: 10.1017/S0266462307051677] [Medline: 17234027]
15. McGee-Lennon M, Wolters M, Clark J. User-centered design of technologies to support care at home. In: Turner KJ, editor. Advances in home care technologies. Amsterdam: IOS Press; 2012:138-161.
16. van Gemert-Pijnen JE, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
17. De Vito Dabbs A, Myers BA, Mc Curry KR, Dunbar-Jacob J, Hawkins RP, Begey A, et al. User-centered design and interactive health technologies for patients. Comput Inform Nurs 2009;27(3):175-183 [FREE Full text] [doi: 10.1097/NCN.0b013e31819f7c7c] [Medline: 19411947]
18. De Rouck S, Jacobs A, Leys M. A methodology for shifting the focus of e-health support design onto user needs: a case in the homecare field. Int J Med Inform 2008 Sep;77(9):589-601. [doi: 10.1016/j.ijmedinf.2007.11.004] [Medline: 18248846]
19. Shah SG, Robinson I, AlShawi S. Developing medical device technologies from users' perspectives: a theoretical framework for involving users in the development process. Int J Technol Assess Health Care 2009 Oct;25(4):514-521. [doi: 10.1017/S0266462309990328] [Medline: 19845981]
20. Poulson D, Ashby M, Richardson S. Userfit: a practical handbook on user-centered design for assistive technology. Brussels: European Commission; 1996.
21. Huis in 't Veld RM, Widya IA, Bults RG, Sandsjö L, Hermens HJ, Vollenbroek-Hutten MM. A scenario guideline for designing new teletreatments: a multidisciplinary approach. J Telemed Telecare 2010;16(6):302-307. [doi: 10.1258/jtt.2010.006003] [Medline: 20798423]
22. Abascal J, Arrue M, Garay N, Thomás J. USERfit tool: a tool to facilitate design for all. In: Universal Access: Theoretical Perspectives, Practice, and Experience. Berlin: Springer; 2002 Presented at: 7th ERCIM International Conference on User Interfaces for All; October 2002; Paris, France.
23. Gulliksen J, Göransson B, Boivie I, Blomkvist S, Persson J, Cajander ASA. Key principles for user-centred systems design. Behaviour & Information Technology 2003 Nov;22(6):397-409. [doi: 10.1080/01449290310001624329]
24. Bridgelal Ram M, Grocott PR, Weir HC. Issues and challenges of involving users in medical device development. Health Expect 2008 Mar;11(1):63-71. [doi: 10.1111/j.1369-7625.2007.00464.x] [Medline: 18275403]

25. Martin JL, Barnett J. Integrating the results of user research into medical device development: insights from a case study. BMC Med Inform Decis Mak 2012;12:74 [FREE Full text] [doi: 10.1186/1472-6947-12-74] [Medline: 22812565]
26. van der Weegen S, Verwey R, Spreeuwenberg M, Tange H, van der Weijden T, de Witte L. The Development of a Mobile Monitoring and Feedback Tool to Stimulate Physical Activity of People With a Chronic Disease in Primary Care: A User-Centered Design. JMIR Mhealth Uhealth 2013 Jul 02;1(2):e8. [doi: 10.2196/mhealth.2526]
27. Verwey R, van der Weegen S, Tange H, Spreeuwenberg M, van der Weijden T, de Witte L. Get moving: the practice nurse is watching you! A case study of the user-centred design process and testing of a web-based coaching system to stimulate the physical activity of chronically ill patients in primary care. Inform Prim Care 2012;20(4):289-298. [Medline: 23890341]
28. Verwey R, van der Weegen S, Spreeuwenberg M, Tange H, van der Weijden T, de Witte L. A pilot study of a tool to stimulate physical activity in patients with COPD or type 2 diabetes in primary care. J Telemed Telecare 2014 Jan;20(1):29-34. [doi: 10.1177/1357633X13519057] [Medline: 24414397]
29. Vermeulen J, Neyens JC, Spreeuwenberg MD, van Rossum E, Sipers W, Habets H, et al. User-centered development and testing of a monitoring system that provides feedback regarding physical functioning to elderly people. Patient Prefer Adherence 2013;7:843-854 [FREE Full text] [doi: 10.2147/PPA.S45897] [Medline: 24039407]
30. Abras C, Maloney-Krichmar D, Preece J. User-centered design. In: Bainbridge W, editor. Encyclopedia of human-computer interaction. Thousand Oaks: Sage Publications; 2004:763-768.
31. Cooper A, Reimann R, Cronin D. About face 3: the essentials of interaction design. Indianapolis, IN: Wiley; 2007.
32. Gasson S. The reality of user-centered design. Journal of End User Computing 1999;11(4):3-13.
33. Goodman-Deane J, Langdon P, Clarkson J. Key influences on the user-centred design process. J of Eng Design 2010 Apr;21(2):345-373. [doi: 10.1080/09544820903364912]
34. Chamberlain S, Sharp H, Maiden N. Towards a framework for integrating agile developmentuser-centred design. In: Abrahamsson P, Marchesi M, Succi G, editors. Extreme programming and agile processes in software engineering. Berlin Heidelberg: Springer; 2006:143-153.
35. Vredenburg K, Mao JY, Smith PW, Carey T. A survey of user-centered design practice. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York: ACM; 2002 Presented at: Conference on Human Factors in Computing Systems; Apr 20-25; Minneapolis, USA p. 471-478.
36. van Limburg M, van Gemert-Pijnen JE, Nijland N, Ossebaard HC, Hendrix RM, Seydel ER. Why business modeling is crucial in the development of eHealth technologies. J Med Internet Res 2011;13(4):e124 [FREE Full text] [doi: 10.2196/jmir.1674] [Medline: 22204896]

Abbreviations

R&D: research and development
UCD: user-centered design

Edited by G Eysenbach; submitted 17.12.13; peer-reviewed by A DeVito Dabbs, S Shah; comments to author 27.01.14; revised version received 17.03.14; accepted 28.04.14; published 19.05.14.

Please cite as:
Vermeulen J, Verwey R, Hochstenbach LMJ, van der Weegen S, Man YP, de Witte LP
Experiences of Multidisciplinary Development Team Members During User-Centered Design of Telecare Products and Services: A Qualitative Study
J Med Internet Res 2014;16(5):e124
URL: http://www.jmir.org/2014/5/e124/
doi:10.2196/jmir.3195
PMID:24840245

Review

eHealth Interventions for HIV Prevention in High-Risk Men Who Have Sex With Men: A Systematic Review

Rebecca Schnall[1], RN, MPH, PhD; Jasmine Travers[1], AGNP-C, RN; Marlene Rojas[1], MPH, MD; Alex Carballo-Diéguez[2], PhD

[1]Columbia University, School of Nursing, New York, NY, United States

[2]Columbia University, HIV Center for Clinical and Behavioral Studies, New York State Psychiatric Institute, New York, NY, United States

Corresponding Author:
Rebecca Schnall, RN, MPH, PhD
Columbia University
School of Nursing
617 West 168th Street
New York, NY,
United States
Phone: 1 212 342 6886
Fax: 1 212 305 6937
Email: rb897@columbia.edu

Abstract

Background: While the human immunodeficiency virus (HIV) incidence rate has remained steady in most groups, the overall incidence of HIV among men who have sex with men (MSM) has been steadily increasing in the United States. eHealth is a platform for health behavior change interventions and provides new opportunities for the delivery of HIV prevention messages.

Objective: The purpose of this systematic review was to examine the use of eHealth interventions for HIV prevention in high-risk MSM.

Methods: We systematically searched PubMed, OVID, ISI Web of Knowledge, Google Scholar, and Google for articles and grey literature reporting the original results of any studies related to HIV prevention in MSM and developed a standard data collection form to extract information on study characteristics and outcome data.

Results: In total, 13 articles met the inclusion criteria, of which five articles targeted HIV testing behaviors and eight focused on decreasing HIV risk behaviors. Interventions included Web-based education modules, text messaging (SMS, short message service), chat rooms, and social networking. The methodological quality of articles ranged from 49.4-94.6%. Wide variation in the interventions meant synthesis of the results using meta-analysis would not be appropriate.

Conclusions: This review shows evidence that eHealth for HIV prevention in high-risk MSM has the potential to be effective in the short term for reducing HIV risk behaviors and increasing testing rates. Given that many of these studies were short term and had other limitations, but showed strong preliminary evidence of improving outcomes, additional work needs to rigorously assess the use of eHealth strategies for HIV prevention in high-risk MSM.

(J Med Internet Res 2014;16(5):e134) doi:10.2196/jmir.3393

KEYWORDS

HIV prevention; eHealth; high-risk MSM; HIV testing; HIV risk behaviors; SMS; Internet

Introduction

Men who have sex with men (MSM) are the population most heavily affected by infection with human immunodeficiency virus (HIV) [1]. The rate of a new HIV diagnosis among MSM is more than 40 times that of women and more than 44 times that of other men [2]. In 2010, male-to-male sex remained the largest HIV transmission category in the United States and the only one associated with an increasing number of HIV/ acquired immunodeficiency syndrome (AIDS) diagnoses [3]. Although MSM represent about 7% of the male population in the United States, they account for 78% of the new HIV infections among males, reinforcing the need for intensive HIV prevention services and testing campaigns [4].

The number of new HIV infections among MSM increased 12% from 2008-2010, with a 22% increase among MSM aged 13-24

years. Notably, young African American MSM account for a disproportionate number of new HIV cases in the United States. There were more new HIV infections (54%) among young African American MSM (aged 13-29 years) than any other racial or ethnic age group of MSM [5], which is nearly twice that of young white MSM and more than twice that of young Hispanic/Latino MSM [6]. Increases in the number of HIV positive individuals in this group suggest that risky sexual behavior has risen despite advances made in testing, prevention, and treatment [5].

While many HIV prevention interventions have been delivered face to face, the emergence of eHealth as a platform for health behavior change provides new opportunities for developing HIV prevention strategies [7]. eHealth is a generic term that applies to an increasingly large number of interventions that are delivered electronically. eHealth can include Web-based tools including videos, games, chat rooms, social networking sites as well as text messaging (SMS, short message service), and email [8]. Across a wide range of diseases and health behaviors, eHealth interventions are successful in promoting changes in behavior, self-efficacy, knowledge, and clinical outcomes. eHealth interventions have been developed to prevent obesity [9,10], treat alcohol abuse [11], promote smoking cessation [12], and encourage nutritious eating [13].

The Internet is an important delivery method for eHealth tools. As access to the Internet increases, Americans' willingness to use the Internet as a source of health information has proliferated, suggesting Web-based interventions are an important modality for health behavior change interventions [14]. Online interventions can be extremely convenient for users as they are accessible from anywhere that there is connection to the Internet and can be used in a private setting, which can also improve accessibility [15]. In a recent systematic review, Guse et al evaluated the impact of digital media-based interventions on the sexual health knowledge, attitudes, and/or behaviors of adolescents aged 13-24 years. Two interventions significantly delayed initiation of sex, and one was successful in encouraging users of a social networking site to remove sex references from their public profile. Seven interventions significantly influenced psychosocial outcomes such as condom self-efficacy and abstinence attitudes, and six studies increased knowledge of HIV, sexually transmitted infections, or pregnancy [16].

A growing number of eHealth HIV prevention interventions have been developed for MSM [17-20]. eHealth interventions are particularly relevant for this high-risk population because of the privacy feature they provide. A user can privately access them without the fear of stigma, which highly affects the MSM community [21]. As the evidence-base on eHealth HIV prevention interventions grows, there is also a need to systematically evaluate the efficacy and feasibility of the existing interventions specific to the MSM population.

Methods

Identification of Studies

We searched articles published from January 2000 to April 2014 in the following electronic databases: PubMed, PsycINFO, Embase, ISI Web of Knowledge, Google Scholar, and Google for grey literature of US and international studies. We visually scanned the reference lists of retrieved documents to identify additional relevant manuscripts. Our search terms included HIV, online, mobile technology, AIDS, technology, electronic health, eHealth, chat room, social networking, mobile applications, mobile health applications, mobile phone, mHealth, text messaging, telemedicine, HIV treatment, PLWH, reminder systems, information systems, Computers, Handheld/ or Cellular Phone/ or mobile applications, HIV/ or HIV.mp or HIV Infections/; Cellular Phone/ or mobile application.mp; HIV/, HIV Infections/ or PLWH.mp, HIV infection, intervention, mobile applications, and mobile HIV applications.

Inclusion Criteria

Included studies had to (1) focus on an eHealth intervention only and could not use eHealth solely as a recruitment or data collection tool, (2) focus on HIV prevention or testing and not on HIV care, (3) be published in English, (4) be published between January 2000 and April 2014, (5) be quasi-experimental or a randomized controlled trial (RCT), (6) have a behavioral outcome measure, and (7) focus on adult MSM. We did not include adolescent studies in our review since a recent systematic review was published [16].

Assessing Study Quality

A quality assessment tool (Table 1) for evaluating HIV prevention interventions was created based on the previously published efficacy criteria developed by the Center for Disease Control and Prevention's HIV/AIDS Prevention Research Branch [22,23]. Papers were scored in each of seven quality domains, and a final total score was calculated as a percentage of possible applicable points. The domains were representativeness, bias and confounding, description of the intervention, outcomes and follow-up, statistical analysis, strength of evidence, and group equivalence. Each of the seven quality domains was given equal weight.

Table 1. HIV prevention intervention quality assessment tool.

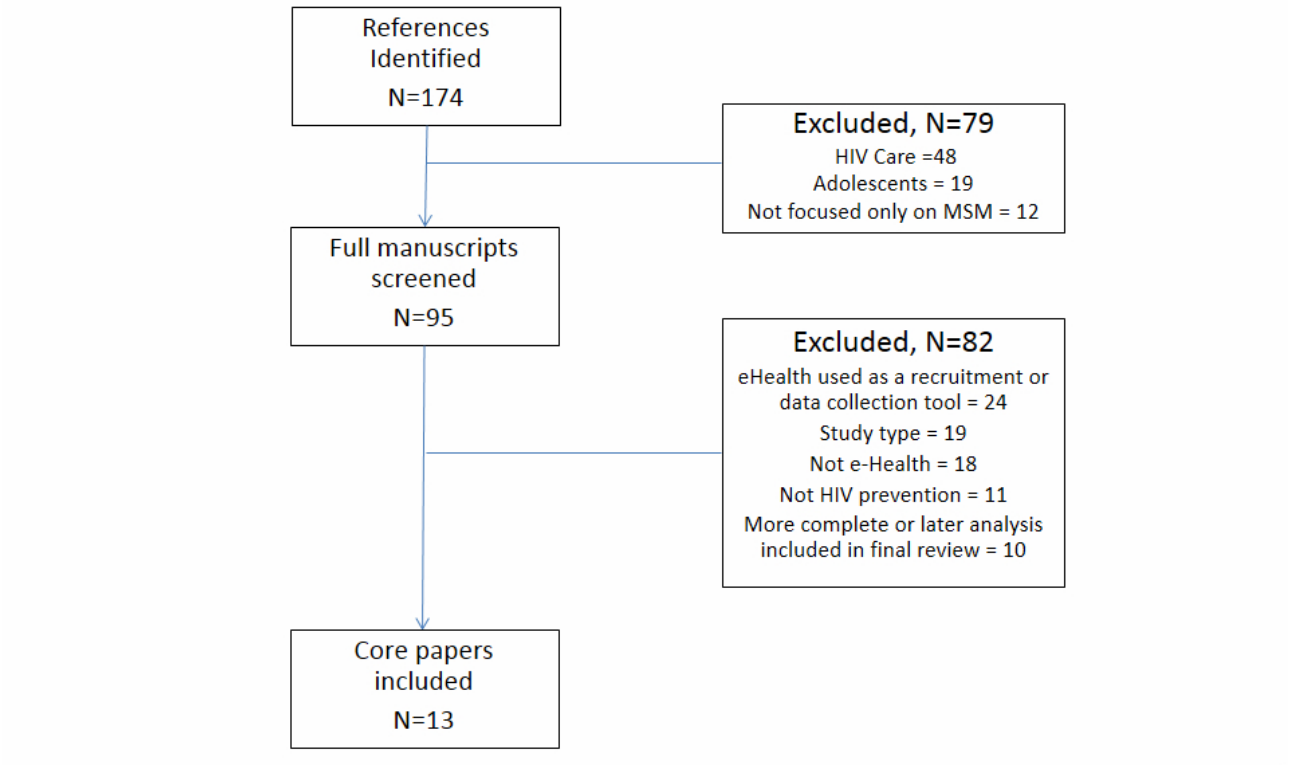
	Completely adequate (%)	Partially adequate (%)	Inadequate, not stated, or impossible to tell (%)
Representativeness	All key characteristics of study population described (50)	Some key characteristics described (25)	Minimal to no description of key characteristics and inclusion/exclusion criteria (0)
	Detailed inclusion/exclusion criteria described (50)	Some description of inclusion/exclusion criteria (25)	
Bias and confounding	Study population corresponded to larger population in all key factors (25)	Sample population differed in some minor factors to larger population (12.5)	Sample population differed in several key factors to larger population (0)
	Equivalent outcome assessment (25)	Minor differences in outcome assessment (12.5)	Major differences in outcome assessment (0)
	Study accounted for confounding interventions with respect to effectiveness of intervention (25)	Study only partially accounted for confounding interventions with respect to effectiveness of intervention (12.5)	Study did not account for confounding interventions with respect to effectiveness of intervention (0)
	Compliance rate >80% (25)	Compliance rate between 80-50% (16.7)	Compliance rate <50% (8.3)
Description of intervention	Protocol could be replicated given description of intervention and /or monitoring (100)	Some minor details excluded from explanation of intervention and/or monitoring (66.7)	No details given in description of intervention and monitoring (0)
		Some major details excluded from explanation of intervention and/or monitoring (33.3)	
Outcomes and follow-up	Outcome assessment procedure clearly defined (50)	Outcome assessment procedure somewhat defined (25)	Outcome assessment procedure not defined (0)
	Groups equivalent in attrition (50)	Some difference in attrition (25)	Major difference in attrition (0)
Statistical analysis	Statistical methods fully described and appropriate (50)	Statistical methods partially described and appropriate (25)	Statistical methods not described or absent (0)
	Tests addressed differences between groups and variability (50)	Tests addressed some differences between groups and variability (25)	Did not address differences between groups and variability (0)
Strength of evidence	Significant positive intervention effects (100)	Significant effect but not in the stated relevant outcome measure (50)	No significant intervention effect (0)
	Positive and statistically significant ($P \leq .05$) intervention effect in ≥1 relevant outcome measure		
Group equivalence	Meets all 4 criteria (100)	Meets 3 criteria (75)	Meets no criteria (0)
	1. Include one or more separate control or comparison study groups.	Meets 2 criteria (50)	
	2. Include clear description of study group comparability.	Meets 1 criteria (25)	
	3. Include clear description of randomization method used or rationale for not using randomization technique in instances when it is not feasible		
	4. Include appropriate statistical controls when equivalence is not achieved		

Data Extraction

Figure 1 summarizes the search results and the outcome of the screening process. The search identified 174 unique papers. These papers were independently appraised by 2 authors with no blinding to the authorship of the papers. Following our appraisal of the abstracts, we excluded articles that were focused on HIV care for persons living with HIV, 19 articles focused on adolescents, and 12 articles that did not solely focus on MSM (eg, women, all men, mixed genders). We reviewed the full manuscript of the remaining 95 articles. Of these, we excluded 10 articles because there was a later published article that was more recent and/or of higher quality. We also excluded 19 other articles because they used a qualitative research design or did not have a behavioral outcome (see Figure 1 for further details of articles that were excluded).

Data were extracted based on objectives, study design, sample size, type and duration of interventions, outcome measures reported, and findings. To further characterize the intervention, we abstracted the theoretical framework used to guide the intervention design, if reported. Data were also abstracted according to country.

Figure 1. Screening process flowchart.



Characteristics of Evidence

A total of 13 articles met the inclusion criteria. The articles were published between 2008 and 2013. Table 2 describes each paper by study design, eHealth strategy, theoretical framework, length of intervention and follow-up, study population, results, and quality rating [24-36]. The total sample in each study ranged from 52 [24] to 3092 men [25]. Nine of the studies were conducted in the United States, and the remaining studies were conducted in Peru (n=1), Australia (n=1), Taiwan (n=1), and Hong Kong (n=1).

Each of the studies had different interventions. Interventions were not clearly described in approximately 46% of the studies. The length of the intervention period ranged from about 15 minutes [25] to 6 months [26-29]. Consistent with inclusion criteria, all studies targeted MSM, with one study [30] focusing on rural MSM and one study [24] focusing on methamphetamine users. Outcome measures included condom use, HIV testing rates, and sexual risk behaviors. eHealth strategies included Web-based videos and education modules (7 studies), text messaging (3 studies), chat room intervention (1 study), and social networking (2 studies). Details of the interventions are described below. A majority (69%) of the interventions were guided by a theoretical framework including the Information-Motivation-Behavioral Skills Model (IMB), the Health Belief Model, Stages of Change, Social Learning Theory, Social Cognition and Developmental Theory, and the Sexual Health Model. The remaining interventions were not guided by theoretical frameworks, or details were not provided in the article. Eight of the studies used an RCT and five of the studies were quasi-experimental. The methodological quality of the available evidence varied, and none of the included studies fulfilled all of the criteria, with quality scores ranging from 49.4-94.6%.

Table 2. Existing studies of eHealth HIV prevention interventions for adult MSM.

Study	Study design	eHealth Strategy	Length of study	Study population	Results	Mean quality score (range)
Blas, 2010 [31]	RCT	Web-based Intervention	Mean of 125.5 days of observation	Intervention (N=239); Comparison (N=220)	Increased HIV testing rates	76.79% (50-100)
Bourne, 2011 [26]	Pre-post test design	SMS reminders	SMS reminders every month for 3-6 months	Intervention (N=714); Control (N=1084)	Increased HIV re-testing rates	51.14 (25-100)
Bowen, 2008 [30]	Pre-Post study	Web-based education modules	Mean 19.39 days (SD 7.33 days)	Rural MSM (N=475)	Decrease high-risk sexual risk behaviors;	89.89 (50-100)
Carpenter, 2010 [32]	RCT	Web-based skills training and motivational intervention	Intervention 1.5-2 h; 3-month follow up	MSM (N=112)	Reduction in high-risk HIV behavior	64.89 (0-100)
Christensen, 2013 [33]	RCT	Virtual Simulation Intervention	3-month follow-up questionnaire	Intervention (N=437) Control (N=484)	Shame reduction; shame reduction as a predictor of UAI[a]	75 (33.3-100)
Hirshfield, 2012 [25]	RCT	Web-based media intervention (prevention videos & webpage)	Baseline survey, Intervention 60 day follow up	Intervention (N=2483) Control (N=609)	More likely to disclose HIV status to partners; less likely to report UAI	89.89 (54.2-100)
Ko, 2013 [27]	Quasi-Experimental, Non-Equivalent control	Web-based peer leader intervention	Baseline survey, 6-month intervention, follow-up survey	Intervention (N=499); Comparison (N=538)	Increased HIV testing, reduced UAI	60.11 (33-100)
Lau, 2008 [28]	RCT	Web-based educational tool	6-month study period	Intervention (N=140); Control (N=140)	Efficacy of the intervention was not supported	65.49 (0-100)
Mustanski, 2013 [34]	RCT	Web-based media intervention	12-wk study period	Intervention (N=50); Control (N=52)	Decrease sexual risk behavior	94.64 (75-100)
Reback, 2012 [24]	Pre-post test design	Text Messaging	2-wk intervention	Meth-using MSM (N=52)	Decreased frequency of methamphetamine use; Decrease high-risk sexual behaviors.	83.93 (50-100)
Rhodes, 2011 [29]	Single-group pretest-post-test design	Chat Rooms	6-month implementation phase; 1-month follow-up	MSM (N=346 [pretest], 315 [posttest])	Increased HIV testing rates	64.89 (25-100)
Rosser, 2010 [35]	RCT	Interactive Website	3-wk intervention	MSM (N=650)	Reduction in risk behavior	49.41 (25-66)
Young, 2013 [36]	RCT	Web Based, Peer leader led groups	12-wk intervention; 12-wk follow-up	112 MSM Intervention N=55 Control N=57	Increased requests for an HIV home test	80.36 (50-100)

[a]UAI: unprotected anal intercourse

Web-Based Videos and Education Modules (k=7)

Two studies used videos for educating high-risk MSM. In a study conducted in Peru, 5-minute videos were created using the Health Belief Model and Stages of Change Theory to encourage MSM to get tested for HIV. The videos incorporated ways to overcome eight reasons why MSM do not get tested for HIV (eg, fear or lack of confidentiality) [31]. In another study, informed by the Social Learning Theory and Social Cognition and Developmental Theory, five study conditions were compared using an RCT and included a (1) dramatic video, (2) documentary video, (3) both dramatic and documentary videos, (4) prevention webpage, and (5) control (received no intervention) [25]. "The Morning After" is a 9-minute video drama [37] that was designed to promote critical thinking about HIV risk and features 3 gay male friends, one of whom thinks he had unprotected sex with an HIV-positive man while intoxicated and seeks advice from friends."*Talking About HIV*" is a 5-minute documentary video created with footage from the documentary, "Meth" [38]. In the 60 days after the intervention, men in the pooled video group were significantly more likely than men in the control group to report full disclosure with their last sexual partner (OR 1.32, 95% CI 1.01-1.74). HIV positive men in this group were also significantly more likely to reduce unprotected anal intercourse (UAI; OR 0.38, 95% CI 0.20-0.67) and serodiscordant UAI (OR 0.53, 95% CI 0.28-0.96) at follow-up. Findings from this study suggest that a brief digital

media intervention can decrease sexual risk behaviors and increase HIV disclosure to potential sexual partners [25].

One study developed and tested multicomponent Internet sites that targeted high-risk sexual behaviors. The intervention, Sexpulse, was a multifaceted Internet intervention that targeted men who use the Internet to seek sex with men and was informed by the Sexual Health Model. Sexpulse was designed by a multidisciplinary team of health professionals, computer scientists, and e-learning specialists and had the following components: a risk assessment tool, an online chat simulation, and virtual peers. Use of the system successfully reduced high-risk sexual behavior in study participants [35]. Carpenter et al developed an Internet site [39] based on the Information Motivation Behavioral Skills theory of HIV risk reduction, which included risk assessment and feedback, motivational exercises, skills training, and education. The format of the material was designed to engage younger MSM, including those from minority groups. Both the intervention and control groups demonstrated reductions in high-risk sexual activity; the intervention group showed greater reductions with the riskiest partners [32].

Keep It Up! (KIU!) was an online, interactive HIV prevention program. The IMB model of HIV risk behavior change was used to guide the development of the KIU! intervention. It has 7 modules completed across 3 sessions that were done at least 24 hours apart and takes about 2 hours to complete. Keep It Up! was designed to be delivered to young MSM upon receiving an HIV negative test result. In an RCT, the participants in the intervention arm had a significantly lower rate of unprotected anal sex acts at the 12-week follow-up [34].

Socially Optimized Learning in Virtual Environments (SOLVE) is a downloadable simulation video game designed to simulate and immerse high-risk young adult MSM in affectively charged risky situations. This intervention was informed by the Theory of Planned Behavior, and Social Cognitive Theory. Christensen et al tested this intervention compared to a wait list control condition in an online RCT. After 3 months, participants in the SOLVE treatment condition reported greater reductions in shame. The direct effect on risky sexual behavior at follow-up was not significant [33].

Finally one study developed and tested a Web-based education module tailored to the information needs of MSM residing in rural areas. There were two 20-minute education sessions that participants watched 6 months apart. Each session consisted of three modules focused on the concepts in the IMB model. Post-intervention behavior change included reduced anal sex and significant increases in condom use [30].

Text Messages or Short Messaging Service and Email Messaging (k=3)

Three studies in our review used text messaging or short messaging service (SMS) as an intervention; two studies used it to increase HIV testing rates and one study to reduce high risk behaviors. The two studies that targeted increasing HIV testing rates were conducted outside of the United States. In one of the SMS studies set in Australia, clinicians sent reminders to patients who had previously come to a sexual health clinic to come back for follow-up testing. SMS reminders increased HIV re-testing rates after 9 months [26]. In the Project Tech Support study, participants received 1-3 social support and health education text messages per day for 2 weeks. The goal of the messages was to reduce methamphetamine use and high-risk sexual behaviors. A total of 400 text messages were developed for this study based on the behavioral change theories of Social Support Theory, the Health Belief Model, and Social Cognitive Theory. Participants reported a significant decrease in methamphetamine use and reductions in high-risk sexual behaviors [24]. In the study conducted in Hong Kong, email messages relating to prevention of STI (sexually transmitted infections) and HIV were sent to participants on a biweekly basis [28]. The contents of the emails covered areas of information and discussion about modes of HIV transmission, correct condom use, HIV testing, “relationships & love”, and the relationship between drugs and sex. The goal of this intervention was to reduce HIV risk-related behaviors; however, there were no significant findings.

Chat Room Intervention (k=1)

One study used a chat room intervention named CyBER/testing, informed by the Natural Helping Theory, in which an interventionist entered the chat room from 9 a.m.-5 p.m., Monday to Friday [29]. Few details about the chat room were included in the study to protect the participants who still use the site. The chat room was designed for social and sexual networking among MSM. Every 30 minutes, the interventionist would post information about HIV testing. More specifically, he answered questions about testing processes and locations, referred chatters to other resources, explained HIV infection, and provided information about resources for those who are seropositive (including medical resources and AIDS drug assistance). He would also respond to chat room members who sent him “instant messages”. The intervention significantly increased self-reported HIV testing among chatters overall.

Social Networking Intervention (k=2)

In the HOPE study, social network sites were used for the delivery of HIV prevention information; 16 peer leaders were randomly assigned to deliver information about HIV (intervention) or general health (control) via Facebook groups for over 12 weeks. Participants randomized to the HIV prevention information group were significantly more likely to request an HIV testing kit than control group participants [36]. There were sparse data on returned tests and follow-up test results indicating that even though this intervention influenced participants’ decision to request an HIV test, it did not necessarily impact actual testing behaviors [36].

In another social networking intervention study, Internet popular opinion leaders (iPOL) were used to disseminate HIV prevention information via popular social networking sites [27]. At the 6-month follow-up after the intervention was conducted, MSM who visited the intervention website were more likely to have been tested for HIV (P<.001) and consistently use condoms during anal sex with online sex partners than those using the control website (34.15% versus 26.19%, P=.004). This study used a non-equivalent quasi-experimental design. There were additional flaws in the study design that limit the evidence of

the use of this intervention for improving HIV prevention behaviors including contamination between study groups and self-report of HIV testing and risk behaviors.

Discussion

Principal Findings

This review of eHealth interventions for HIV prevention among adult MSM has drawn together the evidence base specific to behavioral interventions for MSM and found evidence for eHealth interventions being associated with reductions in high risk behaviors and increases in HIV testing rates. Nonetheless, the studies that demonstrated a decrease in sexual risk behavior had different study designs and outcome measures that make it difficult to synthesize the evidence.

Only one US study in our review solely focused on HIV testing as an outcome measure. Given that the US National HIV/AIDS Strategy has established a goal of increasing the awareness of HIV status in the US population from 79% to 90% by 2015, HIV testing is an important HIV prevention measure. In fact, current recommendations are to repeat HIV testing every 3-6 months for high-risk MSM [40]. Thus, there is a need to develop and test eHealth interventions targeted to improve HIV testing rates in high-risk MSM. Given that most of the eHealth intervention studies conducted abroad targeted HIV testing behaviors, future work in the United States should focus on the lessons learned from those studies.

Bourne et al found that SMS can be used to increase HIV testing rates in high-risk MSM [26]. In a single study by Blas et al [31], the use of an online video-based intervention was shown to increase HIV testing in high-risk MSM. Given the growing HIV epidemic in high-risk MSM in the United States and the need to increase HIV testing, both online video-based interventions and certainly SMS should be employed as strategies to increase HIV testing rates. The use of SMS for improving HIV testing rates was evidenced in studies outside of the United States showing promise for its continued use both abroad and in the United States.

From the results presented above, we can infer that eHealth interventions reduce risky sexual behaviors and increase HIV testing. This review has provided evidence that eHealth interventions have the potential for promoting HIV prevention behaviors in adult MSM. Even so, there are a number of limitations in many of the studies we reviewed. For example, in the study conducted by Reback et al (2012) [24], there was no control group, the intervention group was quite small, and most of the study participants were unemployed, which is not necessarily representative of the MSM population. The study intervention was staff intensive and used two-way pagers that no longer exist, limiting the potential for harnessing this technology for future intervention study.

In another study, Young et al (2013) [36] had planned 7 clusters per study arm but ended up with only 2. Recruitment appeared to have been difficult; they used only Facebook, no sex websites (which could have been more efficient to reach people having high risk behavior). To assess their outcome measure, they used HIV tests that needed to be sent to a laboratory. Participants could have used an HIV home test, which may have reduced some of the access barriers. This review highlights the need for the collection of rigorous data measures for understanding outcomes.

Moreover, there is a need for long-term (12 months) follow-up data after the completion of eHealth HIV prevention interventions. In our review, only 1 study assessed the long-term effects (12 months) of the eHealth intervention and found that it did not have a long-term effect on reducing sexual risk behaviors [35], perhaps because this was not a long-term intervention. Since eHealth interventions appear potentially useful for reducing HIV risk behaviors and increasing HIV testing rates, future research should focus on establishing long-term effectiveness as well as comparing the effectiveness of different interventions.

Limitations

Several limitations of this review should be considered when interpreting the findings. The potential heterogeneity of interventions and outcomes are important to note and make the synthesis of the evidence from these studies challenging. Notably, even though we attempted to be as inclusive as possible, our searches may have excluded relevant studies from this systematic review that did not meet our search word criteria, and/or we excluded conference abstracts that met this review's criteria but were not peer-reviewed articles.

Conclusions

Our results have important implications for the use of eHealth interventions for HIV prevention in MSM. This review demonstrates eHealth interventions appear potentially useful for reducing HIV risk behavior and increasing HIV testing rates. The detailed data across the studies allows us to comprehensively identify and describe elements that are essential to the effectiveness of eHealth interventions for promoting HIV prevention among adult MSM. Given the limitations of many of these studies as well as the potential for eHealth to transform health behaviors, additional work needs to rigorously assess the use of eHealth strategies for HIV prevention in high-risk MSM. Future work is needed that employs these interventions in longer and larger trials and to assess their efficacy in improving outcomes.

Acknowledgments

This publication was supported by a cooperative agreement between Columbia University School of Nursing and the Centers for Disease Control and Prevention (CDC; 1U01PS00371501; PI: R Schnall). RS is supported by the National Center for Advancing Translational Sciences, National Institutes of Health, through Grant No. KL2 TR000081, formerly the National Center for Research

Resources, Grant No. KL2 RR024157. The findings and conclusions in this report do not necessarily represent the views of the National Institute of Health or the CDC.

Conflicts of Interest

None declared.

References

1. Hall HI, Song R, Rhodes P, Prejean J, An Q, Lee LM, HIV Incidence Surveillance Group. Estimation of HIV incidence in the United States. JAMA 2008 Aug 6;300(5):520-529 [FREE Full text] [doi: 10.1001/jama.300.5.520] [Medline: 18677024]
2. Centers for Disease Control and Prevention. CDC Analysis Provides New Look at Disproportionate Impact of HIV and Syphilis Among US Gay and Bisexual Men.: The National Center for HIV/AIDS; 2010. URL: http://www.cdc.gov/nchhstp/newsroom/2010/msmpressrelease.html [accessed 2014-05-18] [WebCite Cache ID 6Pg53lj8P]
3. Centers for Disease Control and Prevention. Today's HIV/AIDS Epidemic. URL: http://www.cdc.gov/nchhstp/newsroom/docs/hivfactsheets/todaysepidemic-508.pdf [accessed 2014-05-18] [WebCite Cache ID 6Pg3vDVHV]
4. Center for Disease Control. 2011. HIV among African Americans URL: http://www.cdc.gov/hiv/topics/aa/ [accessed 2014-05-18] [WebCite Cache ID 6Pg3yFdat]
5. Centers for Disease Control and Prevention (CDC). Vital signs: HIV infection, testing, and risk behaviors among youths - United States. MMWR Morb Mortal Wkly Rep 2012 Nov 30;61(47):971-976 [FREE Full text] [Medline: 23190571]
6. Centers for Disease Control and Prevention (CDC). Trends in HIV/AIDS diagnoses among men who have sex with men--33 states, 2001-2006. MMWR Morb Mortal Wkly Rep 2008 Jun 27;57(25):681-686 [FREE Full text] [Medline: 18583954]
7. Noar SM. eHealth Applications: Promising Strategies for Behavior Change. New York: Routledge; 2012.
8. 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]
9. Glasgow RE, Boles SM, McKay HG, Feil EG, Barrera M. The D-Net diabetes self-management program: long-term implementation, outcomes, and generalization results. Prev Med 2003 Apr;36(4):410-419. [Medline: 12649049]
10. McKay HG, King D, Eakin EG, Seeley JR, Glasgow RE. The diabetes network internet-based physical activity intervention: a randomized pilot study. Diabetes Care 2001 Aug;24(8):1328-1334. [Medline: 11473065]
11. Saitz R, Helmuth ED, Aromaa SE, Guard A, Belanger M, Rosenbloom DL. Web-based screening and brief intervention for the spectrum of alcohol problems. Prev Med 2004 Nov;39(5):969-975. [doi: 10.1016/j.ypmed.2004.04.011] [Medline: 15475031]
12. Civljak MS, Stead LF, Hartmann-Boyce J, Sheikh A, Car J. Internet-based interventions for smoking cessation. Cochrane Database Syst Rev 2013;7:CD007078. [doi: 10.1002/14651858.CD007078.pub4] [Medline: 23839868]
13. Alexander GL, McClure JB, Calvi JH, Divine GW, Stopponi MA, Rolnick SJ, MENU Choices Team. A randomized clinical trial evaluating online interventions to improve fruit and vegetable consumption. Am J Public Health 2010 Feb;100(2):319-326 [FREE Full text] [doi: 10.2105/AJPH.2008.154468] [Medline: 20019315]
14. 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]
15. Binik Y. Sexuality and the internet: Lots of hyp(otheses)—only a little data. Journal of Sex Research 2001 Nov;38(4):281-282. [doi: 10.1080/00224490109552098]
16. Guse K, Levine D, Martins S, Lira A, Gaarde J, Westmorland W, et al. Interventions using new digital media to improve adolescent sexual health: a systematic review. J Adolesc Health 2012 Dec;51(6):535-543. [doi: 10.1016/j.jadohealth.2012.03.014] [Medline: 23174462]
17. Moskowitz DA, Melton D, Owczarzak J. PowerON: the use of instant message counseling and the Internet to facilitate HIV/STD education and prevention. Patient Educ Couns 2009 Oct;77(1):20-26. [doi: 10.1016/j.pec.2009.01.002] [Medline: 19217742]
18. Wei C, Lim SH, Guadamuz TE, Koe S. Virtual Versus Physical Spaces: Which Facilitates Greater HIV Risk Taking Among Men Who Have Sex with Men in East and South-East Asia? AIDS Behav 2013 Sep 29:-. [doi: 10.1007/s10461-013-0628-4] [Medline: 24077974]
19. Williams M, Bowen A, Ei S. An evaluation of the experiences of rural MSM who accessed an online HIV/AIDS health promotion intervention. Health Promot Pract 2010 Jul;11(4):474-482 [FREE Full text] [doi: 10.1177/1524839908324783] [Medline: 19116419]
20. Hightow-Weidman LB, Pike E, Fowler B, Matthews DM, Kibe J, McCoy R, et al. HealthMpowerment.org: feasibility and acceptability of delivering an internet intervention to young Black men who have sex with men. AIDS Care 2012;24(7):910-920 [FREE Full text] [doi: 10.1080/09540121.2011.647677] [Medline: 22272759]
21. Williams ML, Bowen AM, Horvath KJ. The social/sexual environment of gay men residing in a rural frontier state: implications for the development of HIV prevention programs. J Rural Health 2005;21(1):48-55 [FREE Full text] [Medline: 15667009]

22. Lyles CM, Crepaz N, Herbst JH, Kay LS, HIV/AIDS Prevention Research Synthesis Team. Evidence-based HIV behavioral prevention from the perspective of the CDC's HIV/AIDS Prevention Research Synthesis Team. AIDS Educ Prev 2006 Aug;18(4 Suppl A):21-31. [doi: 10.1521/aeap.2006.18.supp.21] [Medline: 16987086]
23. Flores SA, Crepaz N, HIV Prevention Research Synthesis Team. Quality of study methods in individual- and group-level HIV intervention research: critical reporting elements. AIDS Educ Prev 2004 Aug;16(4):341-352. [doi: 10.1521/aeap.16.4.341.40396] [Medline: 15342336]
24. Reback CJ, Grant DL, Fletcher JB, Branson CM, Shoptaw S, Bowers JR, et al. Text messaging reduces HIV risk behaviors among methamphetamine-using men who have sex with men. AIDS Behav 2012 Oct;16(7):1993-2002. [doi: 10.1007/s10461-012-0200-7] [Medline: 22610370]
25. Hirshfield S, Chiasson MA, Joseph H, Scheinmann R, Johnson WD, Remien RH, et al. An online randomized controlled trial evaluating HIV prevention digital media interventions for men who have sex with men. PLoS One 2012;7(10):e46252 [FREE Full text] [doi: 10.1371/journal.pone.0046252] [Medline: 23071551]
26. Bourne C, Knight V, Guy R, Wand H, Lu H, McNulty A. Short message service reminder intervention doubles sexually transmitted infection/HIV re-testing rates among men who have sex with men. Sex Transm Infect 2011 Apr;87(3):229-231. [doi: 10.1136/sti.2010.048397] [Medline: 21296796]
27. Ko NY, Hsieh CH, Wang MC, Lee C, Chen CL, Chung AC, et al. Effects of Internet popular opinion leaders (iPOL) among Internet-using men who have sex with men. J Med Internet Res 2013;15(2):e40 [FREE Full text] [doi: 10.2196/jmir.2264] [Medline: 23439583]
28. Lau JT, Lau M, Cheung A, Tsui HY. A randomized controlled study to evaluate the efficacy of an Internet-based intervention in reducing HIV risk behaviors among men who have sex with men in Hong Kong. AIDS Care 2008 Aug;20(7):820-828. [doi: 10.1080/09540120701694048] [Medline: 18608057]
29. Rhodes SD, Vissman AT, Stowers J, Miller C, McCoy TP, Hergenrather KC, et al. A CBPR partnership increases HIV testing among men who have sex with men (MSM): outcome findings from a pilot test of the CyBER/testing internet intervention. Health Educ Behav 2011 Jun;38(3):311-320 [FREE Full text] [doi: 10.1177/1090198110379572] [Medline: 21393625]
30. Bowen AM, Williams ML, Daniel CM, Clayton S. Internet based HIV prevention research targeting rural MSM: feasibility, acceptability, and preliminary efficacy. J Behav Med 2008 Dec;31(6):463-477 [FREE Full text] [doi: 10.1007/s10865-008-9171-6] [Medline: 18770021]
31. Blas MM, Alva IE, Carcamo CP, Cabello R, Goodreau SM, Kimball AM, et al. Effect of an online video-based intervention to increase HIV testing in men who have sex with men in Peru. PLoS One 2010;5(5):e10448 [FREE Full text] [doi: 10.1371/journal.pone.0010448] [Medline: 20454667]
32. Carpenter KM, Stoner SA, Mikko AN, Dhanak LP, Parsons JT. Efficacy of a web-based intervention to reduce sexual risk in men who have sex with men. AIDS Behav 2010 Jun;14(3):549-557 [FREE Full text] [doi: 10.1007/s10461-009-9578-2] [Medline: 19499321]
33. Christensen JL, Miller LC, Appleby PR, Corsbie-Massay C, Godoy CG, Marsella SC, et al. Reducing shame in a game that predicts HIV risk reduction for young adult MSM: a randomized trial delivered nationally over the Web. J Int AIDS Soc 2013;16(3 Suppl 2):18716 [FREE Full text] [Medline: 24242264]
34. Mustanski B, Garofalo R, Monahan C, Gratzer B, Andrews R. Feasibility, acceptability, and preliminary efficacy of an online HIV prevention program for diverse young men who have sex with men: the keep it up! intervention. AIDS Behav 2013 Nov;17(9):2999-3012. [doi: 10.1007/s10461-013-0507-z] [Medline: 23673793]
35. Rosser BR, Oakes JM, Konstan J, Hooper S, Horvath KJ, Danilenko GP, et al. Reducing HIV risk behavior of men who have sex with men through persuasive computing: results of the Men's INTernet Study-II. AIDS 2010 Aug 24;24(13):2099-2107 [FREE Full text] [doi: 10.1097/QAD.0b013e32833c4ac7] [Medline: 20601853]
36. Young SD, Cumberland WG, Lee SJ, Jaganath D, Szekeres G, Coates T. Social networking technologies as an emerging tool for HIV prevention: a cluster randomized trial. Ann Intern Med 2013 Sep 3;159(5):318-324. [doi: 10.7326/0003-4819-159-5-201309030-00005] [Medline: 24026317]
37. HIV Big Deal. URL: https://www.hivbigdeal.org [accessed 2014-05-22] [WebCite Cache ID 6PlW8a19H]
38. Ahlberg T. Meth. United States: Cinema Libre Distribution; 2006. URL: http://www.imdb.com/title/tt0783550/ [accessed 2014-05-22] [WebCite Cache ID 6PlVpnKe6]
39. Carpenter K. HotandSafeM4M. URL: http://elearning.talariainc.com/buildcontent.aspx?tut=132&pagekey=17401&cbreceipt=0
40. Branson B, Handsfield HH, Lampe MA, Janssen RS, Taylor AW, Lyss SB, Centers for Disease Control and Prevention (CDC). Revised recommendations for HIV testing of adults, adolescents, and pregnant women in health-care settings. MMWR Recomm Rep 2006 Sep 22;55(RR-14):1-17; quiz CE1 [FREE Full text] [Medline: 16988643]

Abbreviations

AIDS: acquired immunodeficiency syndrome
CDC: Centers for Disease Control and Prevention

HIV: human immunodeficiency virus
IMB: Information-Motivation-Behavioral Skills Model
MSM: men who have sex with men
RCT: randomized controlled trial
SMS: short message service
UAI: unprotected anal intercourse

Edited by G Eysenbach; submitted 12.03.14; peer-reviewed by S Hirshfield; comments to author 03.04.14; revised version received 25.04.14; accepted 13.05.14; published 26.05.14.

Please cite as:
Schnall R, Travers J, Rojas M, Carballo-Diéguez A
eHealth Interventions for HIV Prevention in High-Risk Men Who Have Sex With Men: A Systematic Review
J Med Internet Res 2014;16(5):e134
URL: http://www.jmir.org/2014/5/e134/
doi:10.2196/jmir.3393
PMID:24862459

Original Paper

Cognitive Factors of Using Health Apps: Systematic Analysis of Relationships Among Health Consciousness, Health Information Orientation, eHealth Literacy, and Health App Use Efficacy

Jaehee Cho[1], PhD; Dongjin Park[2*], PhD; H Erin Lee[3*], PhD

[1]School of Mass Communication, Chung-Ang University, Seoul, Republic of Korea

[2]School of Communication, Hallym University, Chuncheon, Gangwon-do, Republic of Korea

[3]Media Communication Division, Hankuk University of Foreign Studies, Seoul, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Dongjin Park, PhD
School of Communication
Hallym University
#10222 Dasan Building
39 Hallymdaehak-gil
Chuncheon, Gangwon-do,
Republic of Korea
Phone: 82 33 248 1929
Fax: 82 33 256 3424
Email: dongjinpark@hallym.ac.kr

Abstract

Background: Interest in smartphone health apps has been increasing recently. However, we have little understanding of the cognitive and motivational factors that influence the extent of health-app use.

Objective: This study aimed to examine the effects of four cognitive factors—health consciousness, health information orientation, eHealth literacy, and health-app use efficacy—on the extent of health-app use. It also explored the influence of two different use patterns—information and information-behavior use of health apps—with regard to the relationships among the main study variables.

Methods: We collected and analyzed 765 surveys in South Korea. According to the results, there was a negligible gender difference: males (50.6%, 387/765) and females (49.4%, 378/765). All participants were adults whose ages ranged from 19 to 59. In order to test the proposed hypotheses, we used a path analysis as a specific form of structural equation modeling.

Results: Through a path analysis, we discovered that individuals' health consciousness had a direct effect on their use of health apps. However, unlike the initial expectations, the effects of health information orientation and eHealth literacy on health-app use were mediated by health-app use efficacy.

Conclusions: The results from the path analysis addressed a significant direct effect of health consciousness as well as strong mediating effects of health-app use efficacy. These findings contribute to widening our comprehension of the new, digital dimensions of health management, particularly those revolving around mobile technology.

(J Med Internet Res 2014;16(5):e125) doi:10.2196/jmir.3283

KEYWORDS

health apps; health consciousness; health information orientation; eHealth literacy; health app use efficacy

Introduction

Background

Recently, along with the notable development of mobile communication devices, mobile health, known as mHealth, has become one of the hottest issues in the disciplines of medical science, nursing, and health communication [1-3]. The main factors leading this mobile health boom have been the high penetration of Internet access, particularly expanding Wi-Fi-services, continuous improvement of mobile supporting

systems, and increased use of smartphones. For instance, 91% of US adults now own cellular phones [4], and 53% of cell phone users (45% of all US adults) own a smartphone [5]. Supported by these advanced mobile service and technological developments, people have come to actively seek health information through the Internet using their mobile devices [5]. For example, 52% of smartphone users seek health information through their smart devices [5].

In addition to such information-seeking behaviors through smart devices, there has been a notable increase of health apps available on such devices [3,5-12]. According to Kamerow [13], there are approximately 100,000 health-related apps for smartphones; by 2015, about 500 million owners of smartphones throughout the world will use health apps. In 2012, while 84% of smartphone users had downloaded at least one app, 19% of them had used a health app for the purpose of tracking and managing their health [5]. According to the authors, the adoption of health apps was stronger among females, young populations, and people with higher income. These apps provide users with health-related services, such as medical information, blood pressure checks, female health checks, and so on. With regard to the specific functions of health apps, there are three dominant areas: exercise, diet, and weight management [5]. For example, 38% of health-app users use a health app in order to track their exercise [5].

As noted above, the fact that there has been a great increase of health apps on smart devices can hardly be denied [3,5-12]. Thus, recently, previous studies in the areas of medical science and informatics have intensively investigated the effectiveness of specific apps on smartphones [3,5-12]. In particular, they have focused on scrutinizing those functions and features of smartphone apps that are specialized for particular health conditions (eg, obesity) or diseases (eg, diabetes). For example, Morris et al evaluated health applications on cellular phones, particularly for emotional self-awareness [9]. Kirwan et al and Frøisland et al paid significant attention to smartphone apps for Type 1 diabetes [11-12]. These studies meaningfully contributed to widening our understanding of the effectiveness of interventions of particular smartphone apps.

Nevertheless, we have little understanding of the general cognitive motivators that trigger people's use of health apps, which are relevant to individuals' personal psychological conditions. Except for Lim et al's study [14], there has been little empirical research on the cognitive motivational factors of health-app use in general. Without proper comprehension of the motivational and cognitive factors of adopting and using health apps, it would be difficult to fully understand individuals' use of such apps. Hence, this present study aimed at exploring which cognitive factors would lead people to use health apps among smartphone owners in South Korea. This particular country is adequate for the present study due to its high Internet penetration rate as well as its high distribution rate of smartphones. According to the 2013 Internet Use Report [15], the Internet use rate among Korean adults was approximately 86.2% in 2013, and is continuously growing. Moreover, the Internet use rate among young people in their twenties and thirties was approximately 99.7%. Furthermore, according to Google Korea's marketing research on smartphones, smartphone penetration in South Korea reached 73% in July 2013, placing Korea at number one [16]. Notably, approximately 92% of the younger generation in their twenties through thirties own smart devices in Korea [15]. Overall, the findings from this study will help scholars and practitioners widen their comprehension of health-app use.

Theoretical Backgrounds and Hypothesis Building

Although there exist numerous cognitive factors that can potentially stimulate people to use health apps, we paid considerable attention to the following four main factors: (1) health consciousness, (2) health information orientation, (3) eHealth literacy, and (4) health-app use efficacy. We selected these four factors by considering primarily the general functions of health apps. First, because the fundamental function of health apps is to manage one's own health conditions [5], health consciousness is inherently related to health apps. Next, people use health apps in order to seek health information as well as to monitor their health conditions rather than to gain actual physical aid. This implies that health-app use is more relevant to health information-seeking behaviors. Thus, we focused mainly on health information orientation [17]. Third, it must be considered that such health information from health apps requires the competence of users to accurately comprehend the information accessed; this is known as the literacy of health information. Accordingly, paying attention to individuals' abilities to decipher the meaning of Internet health information (eHealth literacy [17]), this study examined the role of eHealth literacy in health-app use. Furthermore, with regard to eHealth literacy, cognitive differences exist in individuals' abilities to find and understand adequate health information [18-19]. This result implies the potential role of health-app use efficacy in mediating the relationship between eHealth literacy and health-app use. With this reasoning, we scrutinized the potential direct and indirect effects of the four cognitive factors on individuals' actual use of health apps. This section will elaborate on each of these factors and propose multiple hypotheses for the study.

First, health consciousness basically refers to the extent to which a person takes care of his/her own health [17,19]. People with higher levels of health consciousness are more likely to have healthy habits, spend more time on exercise and healthy activities, actively gather health information from various sources, and avoid unhealthy situations [17,19]. In particular, such people are interested in seeking a diverse range of health information in order to gather more accurate information [20]. Moreover, previous research has demonstrated that health consciousness positively influences people's information-seeking behaviors on the Internet [17]. Considering this influential role of health consciousness in health information-seeking behaviors, it is quite reasonable to expect that the more conscious a person is of his/her own health, the more actively she/he will use health apps. Based on this argument, this study established the following hypothesis (H1): Health consciousness will be positively associated with the extent of health-app use.

Health orientation is related to an individual's proactive behaviors of taking care of his/her health condition [21].

Moorman and Matulich [22] defined health orientation as "a goal-directed arousal to engage in preventive health behaviors". Specifically relevant to health-information seeking behaviors, Dutta-Bergman conceptualized health information orientation as "the extent to which the individual is willing to look for health information" [17]. More specifically, people with higher levels of health information orientation are more likely to gather health information from various sources. Moreover, Basu and Dutta stated that, "a high level of health-information orientation suggests the willingness to look for issues related to health and to find out ways to educate oneself about these issues, including the consumption of those communication channels that serve as potential sources of information regarding the issue" [19]. Considering the consumption of various channels for health information, health information orientation may be closely related to a person's use of health apps as useful tools for seeking health information. Hence, this study established the following hypothesis (H2): Health information orientation will be positively associated with the extent of health-app use.

Previous research on health information-seeking behaviors has constantly argued the importance of literacy regarding health information from online sources, often referred to as eHealth literacy [5,18,23]. This is because the acquisition of more information does not necessarily mean better information. According to Norman and Skinner, eHealth literacy can be defined as "the ability to seek, find, understand, and appraise health information from electronic sources and apply the knowledge gained to addressing or solving a health problem" [18]. Based on this definition of eHealth literacy, it is comprehensible that when a person can better seek and understand online health information, she/he may be more motivated to use health apps as electronic sources. Thus, this study established the following hypothesis (H3): eHealth literacy will be positively associated with the extent of health-app use.

Related to eHealth literacy, we must consider that individuals have different levels of ability in using health apps. To better understand this notion, the concept of health information efficacy is useful. According to Basu and Dutta-Bergman, health information efficacy basically refers to "the perception of access to or the availability of health information resources" [19]. Furthermore, paying more attention to behavioral aspects, Yun and Park [24] proposed the concept of Internet health information use efficacy that is reliant on the concept of Bandura's self-efficacy [25]. Here, self-efficacy is conceptualized as a person's ability to achieve a directed goal. Thus, Internet health information use efficacy refers to the individual's cognitive ability to strategically seek the necessary information by selectively using certain communicative channels. Based on these arguments regarding Internet health information use efficacy, this current study proposed the concept of health-app use efficacy, which is referred to as the cognitive ability to use health apps in order to access and seek health information.

Here, it is meaningful to focus on the potential relationship between eHealth literacy and health-app use efficacy. As stated above, eHealth literacy is closely related to an individual's cognition of his/her own ability to seek and understand online health information [18,26]. Such self-efficacy related to online behaviors can be significantly associated with the use of mobile tools with online functions—more specifically, health apps on smart devices in this study. In other words, it is plausible that a person with higher levels of eHealth literacy is more likely to perceive that she/he has a better ability to use health apps. This belief implies the positive effect of eHealth literacy on health-app use efficacy. Ultimately, it also depicts that health-app use efficacy mediates the relationship between eHealth literacy and the extent of health-app use. Consequently, this study tested the following hypothesis (H4): Health-app use efficacy will positively mediate the relationship between eHealth literacy and the extent of health-app use.

For a more thorough analysis of the relationships among the main study variables, we differentiated between two different types of health-app users based on the nature of the health apps they use. In order to do this, we must first differentiate between two types of health-app uses: information-oriented use and behavior-oriented use. Information-oriented use refers to searching for health information (eg, symptoms, medication, preventive care) on apps. Behavior-oriented use involves active monitoring, recording, and management of health conditions through apps. Some examples of behavior-oriented use are using the app, "Diabetes in Check", to record and monitor one's daily insulin levels and calorie intake, or using the app, "Runkeeper", to keep track of one's daily fitness routine and history. Recognizing the different functional uses of health apps, the two types of health-app users that we identified are as follows: (1) information-oriented users, or single-purpose users, and (2) information-behavior users, or dual-purpose users. Ultimately, this study explored how the relationships among the five study variables would differ between these two groups of users. For this process, the following research question (RQ1) was explored: How will the relationships among the five study variables differ between single-purpose (information-oriented) users and dual-purpose (information-behavior) users?

Methods

Participants

For this present study, we used a subset of data collected for a larger research project, which examined Koreans' general use of media for health information. The data were collected through an online survey administered by a Korean professional research company, well-known for managing the largest sampling pool in Korea. The sample for the larger project was chosen through a proportionate stratified sampling method, considering gender, age, and residential area. All survey participants were informed of the overall study goals and procedures. Only those who agreed to participate in the online survey were given access to the survey.

The questionnaire, conducted in Korean, included a question that asked the participants to report the different types of media they used in order to search for health information. Only data from those participants that marked the item of "mobile health apps" for this particular question were included in this current study. In other words, all participants included in the current study used health apps for information-oriented purposes.

Through this process, we were able to obtain a total of 765 surveys. There was a negligible difference in terms of gender composition: 50.6% (387/765) were male and 49.4% (378/765) were female. In terms of residential area, participants were from metropolitan areas (40.0%, 306/765), middle-size cities (36.3%, 278/765), and rural areas (23.7%, 181/765). All participants were adults whose ages ranged from 19 to 59 years; the average age was 37.1 years. In terms of educational attainment, the majority of participants held either a college degree (67.5%, 516/765) or a high school degree (22.7%, 249/765). About 9% (69/765) of the participants had graduate degrees (eg, MA or PhD).

In order to identify the different types of health-app users, those participants within our sample who further indicated the use of health apps for behavior-oriented purposes (use of health apps to monitor and manage their health conditions, such as blood pressure, blood sugar, history of exercise, etc) were categorized as information-behavior, dual-purpose users. The remaining participants were categorized as information-oriented, single-purpose users. A slightly larger portion of the survey participants (55.3%, 423/765) engaged in both information-oriented and behavior-oriented use, compared to those who engaged only in information-oriented use (44.7%, 342/765). In terms of the gender composition of the two groups, it was observed that, while there were slightly more female participants in the information-oriented use group (50.6%), there were slightly more male participants in the information-behavior use group (51.5%, 387/765). However, this gender difference was negligible for both groups. The average age of information-oriented users was slightly older (mean 38.9, SD 10.6) than that of information-behavior users (mean 35.7, SD 10.6).

Instruments

Overview

All measures, except for the extent of health-app use, were constructed as 5-point Likert-type scales (eg, 1=strongly disagree, 5=strongly agree). Reliability tests for the four composite measures of this study reached acceptable Cronbach alphas (higher than .70).

Extent of Health-App Use

The extent of health-app use was conceptualized as the intensity of using health apps and was measured through a single item measured on a 6-point scale. Specifically, the participants were asked to answer the following question, “In the last month, how much time did you spend using health apps?” (1=less than 1 hour, 2=1-2 hours, 3=2-4 hours, 4=4-6 hours, 5=6-10 hours, and 6=more than 10 hours). As the extent of health-app use is a unidimensional factor, the use of a single-item measure is quite acceptable [27,28]. Moreover, in place of using the typical Likert-type options, such as a 5-point scale with options such as “much”, “very much”, and “little”, we created and used categories composed of six points, corresponding to the amount of time spent on health apps, in order to obtain a more objective and reliable measurement of the concept. According to previous research [29-31], this type of data format can be used for common parametric tests. Furthermore, following the guidelines from Kline [29] and Lee and Lim [32], the bootstrapping analysis was applied for the path analysis in order to eliminate any standard errors from the non-normal distribution.

Health Consciousness

This factor was measured through Dutta-Bergman’s scale [17], which measured the health-consciousness attitude through five items. Due to a low factor loading score (smaller than .50), one item was removed from further analysis. Examples of the items are as follows: “I am doing relatively well in taking care of my health” and “My health depends on how well I take care of myself”. The reliability score for this measurement was acceptable (mean 3.16, SD 0.64, N=765, alpha=.84).

Health Information Orientation

In order to measure health information orientation, we used Dutta-Bergman’s original scale [17] composed of eight items. In the process of the factor analysis, one item was removed from further analysis due to its low factor loading score. Example of the items are as follows: “To be and stay healthy, it is critical to be informed about health issues” and “When I take medicine, I try to get as much information as possible about its benefits and side effects”. This factor had an acceptable Cronbach alpha score (mean 3.48, SD 0.55, N=765, alpha=.86).

eHealth Literacy

In order to measure eHealth literacy, we used four items from Norman and Skinner’s scale [17]. Examples of these items are as follows: “I know how to find useful health information through the Internet” and “I have the skills I need to evaluate the health resources I find on the Internet”. The reliability score for this measurement was acceptable (mean 3.23, SD 0.59, N=765, alpha=.85).

Health-App Use Efficacy

In order to measure this variable, we reworded four items from Compeau and Higgins’ scale for computer self-efficacy [33]. Examples of those four items are as follows: “It is easy to learn how to use health apps on my smartphone” and “I can evaluate well the quality of health apps on my smartphone”. This factor also had an acceptable Cronbach alpha score (mean 3.23, SD 0.64, N=765, alpha=.87).

Results

Descriptive Statistics

Before conducting the path analysis, we analyzed the descriptive statistics for the five main variables. Through a series of independent samples *t* tests, one-way analysis of variance (ANOVA), and a bivariate correlation analysis, we checked for any differences in the variables in terms of gender, age, education level, and use patterns. First, we found significant gender differences in health consciousness (M_{male}=3.30, M_{female}=3.03, t=6.06, P<.001) and eHealth literacy (M_{male}=3.31, M_{female}=3.15, t=3.83, P<.001). Male participants reported higher scores for these two variables. Next, while age was positively correlated with health information orientation (r=.157, P<.001), it was negatively correlated with health-app use efficacy (r=−.136, P<.001) and the extent of health-app use (r=−.107,

P<.001). Third, the results from ANOVAs reported significant educational differences in health consciousness ($F_{2,762}$=5.20, P=.006) and eHealth literacy ($F_{2,762}$=1.86, P=.019). The post-hoc tests for these two ANOVAs indicated that people with higher educational backgrounds reported higher levels of health consciousness and health-app use efficacy. Last, the results from the independent samples t test indicated that, except for health consciousness, the levels of all other four variables were significantly higher among information-behavior users compared to information-oriented users (see Table 1).

Table 1. Results for independent samples t test between information-oriented and information-behavior users.

Variable	Info-oriented mean (SD)	Info-behavior	t value	P value
Health consciousness (HC)	3.12 (0.64)	3.20 (0.63)	−1.77	.08
Health information orientation (HIO)	3.38 (0.55)	3.60 (0.55)	−4.44	<.001
eHealth literacy (eHL)	3.12 (0.56)	3.32 (0.60)	−4.87	<.001
Health-app use efficacy (HAUE)	3.12 (0.68)	3.40 (0.67)	−5.85	<.001
Extent of health-app use (HAU)	2.84 (1.34)	3.43 (1.51)	−5.59	<.001

Hypotheses Tests

For testing the multiple hypotheses, we developed a path model composed of five paths. In order to test these hypotheses, we conducted a path analysis using AMOS 21 (SPSS software). Further, in order to minimize the standard errors from the non-normal distribution, we followed guidelines from Kline [29] and Lee and Lim [32] and conducted a bootstrapping analysis using a sub-sample of 200 from our study sample. Therefore, the P value for each path was calculated through a bias-corrected percentile method. We checked both the comparative and absolute fit indices in order to evaluate the goodness-of-fit of the proposed path model: comparative fit index (CFI; higher than .90), incremental fit index (IFI; higher than .90), and standardized root-mean squared residual (SRMR; lower than .10). Although the results from the path analysis of the initial model (see Figure 1) presented acceptable model fits (χ^2_2= 27.5, CFI=.95, IFI=.95, SRMR=.04), the modification indices indicated the necessity to add a path from health information orientation to health-app use efficacy. To develop the final model, we removed two insignificant paths and added one path (see Figure 2). As a result, the final model illustrated much better model fits (χ^2_3=1.02, CFI=1.0, IFI=1.0, SRMR=.007). Comparing the initial model to the final model, the chi-square largely and significantly decreased by 26.4 as the degree of freedom increased by one unit. H1 hypothesized a positive association between health consciousness and the extent of health-app use. Fully supporting H1, health consciousness positively and strongly impacted the use of health apps (beta=.286, P=.012).

H2 and H3 focused on the roles of health information orientation and eHealth literacy in directly influencing the extent of health-app use. With regard to these two hypotheses, the results from the path analysis indicated that neither health information orientation (beta=.08, P=.38) nor eHealth literacy (beta=−.09, P=.508) had a direct effect on the extent of health-app use (see Figure 1). These results indicate that H2 and H3 were rejected. However, as the final path model (Figure 2) indicates, health information orientation strongly impacted health-app use efficacy (beta=.220, P=.011). This reveals the indirect effect of health information orientation on the actual use of health apps. Therefore, in order to test the role of health-app use efficacy in mediating the relationship between health information orientation and the extent of health-app use, we used Sobel's test. The test result found a significant mediating effect of health-app use efficacy (Sobel's statistic=2.45, P=.014).

Next, paying attention to the influential cognitive role of self-efficacy in individuals' actual behaviors, we focused on health-app use efficacy. In this study, we hypothesized that health-app use efficacy would positively mediate the relationship between eHealth literacy and extent of health-app use (H4). The results from the path analysis indicated that eHealth literacy strongly and positively affected health-app use efficacy (beta=.39, P=.005), which ultimately impacted the extent of health-app use (beta=.233, P=.023). Additionally, in order to test this mediating effect of health-app use efficacy, we used Sobel's test. The result from Sobel's test fully supported the mediating effect of health-app use efficacy (Sobel's statistic=2.67, P=.007), thereby fully supporting H4.

Last, through RQ1, this study explored how the relationships among study variables would differ across the two groups of health-app users—information-oriented users vs information-behavior users. For this exploration, we conducted a multi-group structural equation modelling (SEM) for the final model (see Figure 2) and compared four pairs of regression coefficients for the two groups. Table 2 shows the results from the statistical comparison. The results indicated that only the path from health-app use efficacy to the extent of health-app use was statistically different between the two groups (Z=−2.14, P=.03). Specifically, while the direct effect of health-app use efficacy on the extent of health-app use was statistically significant among information-behavior users (beta=.319, P=.008), such effect was not significant among information-oriented users (beta=−.045, P=.734).

Table 2. Comparisons of regression coefficients between information-oriented users and information-behavior users.

Path	Information-oriented users, β	Information-behavior users, β	*Z* score	*P* value
HC[a] → HAU[e]	.193	.315	−.73	>.10
HIO[b] → HAUE[d]	.189	.221	−.31	>.10
eHL[c] → HAUE	.352	.42	−.77	>.10
HAUE → HAU	−.045	.319	−2.14	.03

[a]HC: health consciousness

[b]HIO: health information orientation

[c]eHL: eHealth literacy

[d]HAUE: health-app use efficacy

[e]HAU: extent of health-app use

Figure 1. Initial path model of main study variables with entire sample. HC: Health Consciousness; HIO: Health Information Orientation; eHL: eHealth Literacy; HAUE: Health-App Use Efficacy; HAU: Extent of Health-App Use; e1: Standard Error for HAUE; e2: Standard Error for HAU.

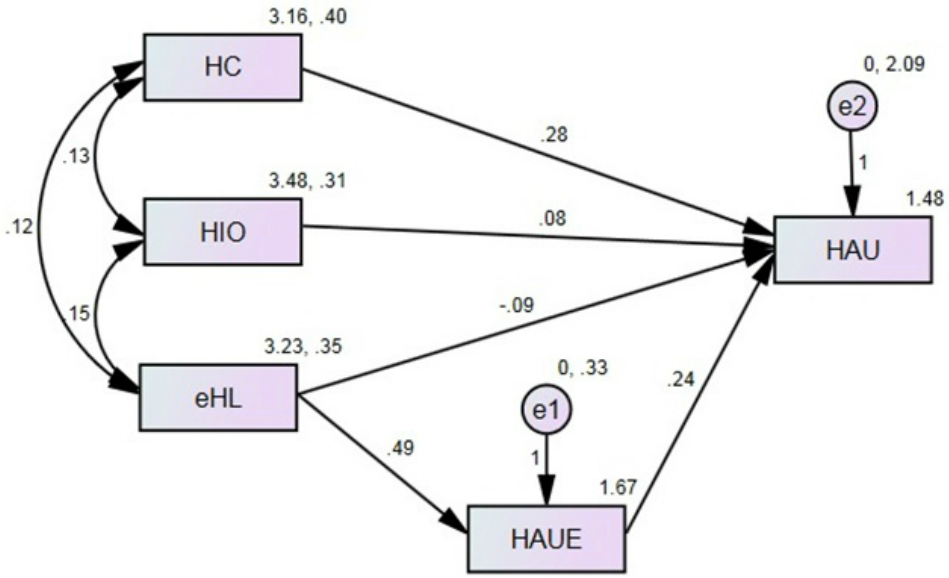
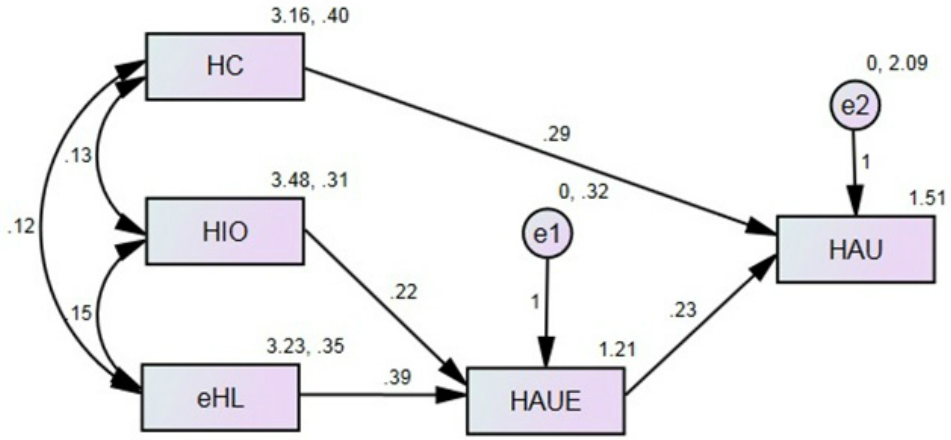


Figure 2. Final path model of main study variables with entire sample. HC: Health Consciousness; HIO: Health Information Orientation; eHL: eHealth Literacy; HAUE: Health-App Use Efficacy; HAU: Extent of Health-App Use; e1: Standard Error for HAUE; e2: Standard Error for HAU.



Discussion

Principal Findings

Considering the lack of studies on the motivational factors of health-app use, the main goal of this study was to explore how cognitive factors would motivate individuals' use of health apps. In particular, paying attention to two different types of health-app use—information-oriented (single-purpose) use and information-behavior (dual-purpose) use—we focused on four cognitive factors: health consciousness, health information orientation, eHealth literacy, and health-app use efficacy. Before conducting the main path analysis, we checked for differences in the five main variables across gender, age, education, and patterns of health-app use. Consistent with the previous research on health apps [5], younger participants reported higher scores for both health-app use efficacy and extent of health-app use; further, participants with higher educational backgrounds tended to be more conscious of their health and have higher levels of eHealth literacy. Moreover, this study found that, compared to women, men also reported higher levels of health consciousness and eHealth literacy. These results support the findings of the previous research. Another notable finding is that information-behavior users reported higher scores for most variables compared to information-oriented users.

In addition to these descriptive findings, a number of meaningful findings were observed through a path analysis. First, supporting the findings of previous research regarding the positive functions of health consciousness [17], health consciousness in this study was also found to be positively and significantly associated with individuals' use of health apps. That is, individuals more interested in taking care of themselves are more likely to use health apps than individuals less conscious of their health. This finding reaffirms the existing knowledge that understands health consciousness to be one of the most dominant factors guiding the adoption of health technologies.

Next, it is noteworthy that, unlike the initial predictions, there was no direct effect of health information orientation and eHealth literacy on the extent of health-app use. Rather, the effects of these two factors were mediated by health-app use efficacy. This indicates the significant roles of health-app use efficacy. In general, people who are more oriented toward actively seeking health information and having a better understanding of online health information tend to be more efficient in using health apps as well as allocating more time for health apps. However, as the results from the multi-group SEM show, the effect of health-app use efficacy on the extent of health-app use was statistically significant only among information-behavior users. This may be due to the displacement of media for health information [34,35]. Information-oriented users may irregularly and occasionally use health apps only at those times that they are in need of certain types of health information. On the other hand, information-behavior users also tend to occasionally seek health information, but manage their health on a regular basis. Here, it must be considered that general health information can be obtained through various online sources. This implies that health apps for general health information can be more easily displaced by these convenient alternatives. However, considering their habitual use of health apps for health management, information-behavior users may need to invest in additional resources in order to seek and routinize alternative media. Moreover, when a person is efficiently using a certain health app on a regular basis, she/he will be more inclined to continue using the app and more reluctant to displace it. This particular finding addresses the key roles of health-app use patterns for determining the intensity/extent of health-app use and, further, theoretically contributes to widening our understanding of the behavioral aspects of health-app use.

Furthermore, this significant role of health-app use efficacy suggests the following practical implication. Health-app use efficacy is conceptually reliant on Bandura's [25] concept of self-efficacy. Moreover, with regard to technology use, self-efficacy is often related to the perceived ease of using certain technologies. Consequently, when an individual perceives higher ease of using health apps, she/he may feel higher health-app use efficacy. This addresses the importance of creating health apps that allow higher levels of ease and convenience in use. Although popular health apps provide users with detailed, useful information, many of them require users to complete multiple steps in order to access such information. For example, in order to obtain information about one's daily calorie intake through the smartphone app, "Lose It", users must first complete several steps and provide many details about their meals (eg, exact categories and amounts of each component of their breakfast, lunch, dinner, and snacks). Although this app provides users with accurate information, the repetitive and complex process requires users to invest a great amount of time and mental energy. This may negatively impact users' health-app use efficacy, ultimately affecting their willingness to use the app. Therefore, there is a need for practitioners to work on the simplification and reduction of algorithms in constructing health app processes.

Limitations

Overall, these findings will serve as helpful empirical and theoretical foundations for future research on health apps. Moreover, they may guide practitioners in developing more realistic and strategic plans to enhance health-app consumption. Nevertheless, considering the limitations of this present study, the following points are recommended for future research. First, as stated above, people use health apps for different reasons—to exercise, to lose weight, to check blood sugar, to track period cycles, and so forth. Based on the above uses and the gratification theory [36], the different purposes for using health apps may be related to motivational factors. Although this present study focused on two different general patterns of health-app use—information-oriented use and information-behavior use—future research will benefit from a more thorough exploration of the more diverse range of functions that health apps have, particularly the functions of those that focus on specific types of health conditions and needs (eg, apps for Type 2 diabetes, Alzheimer's, pregnancy, fitness). Moreover, it is possible that the time spent on health apps may be determined by the specific functions afforded by the health apps. For instance, the use of health apps to check for blood sugar levels will require much shorter amounts of time than the

use of fitness apps that track how many miles one has been running; although it is possible that the former type of apps may be used on more frequent levels. Consequently, for future research, it is recommended to grant attention to the multiple aspects of health-app use, particularly with regard to the extent and frequency of use, based on a more detailed identification of the specific features and functions of various apps.

Next, another limitation of this current study is the relatively large portion of participants with college degrees. Although it has been often observed that individuals with higher levels of education tend to more actively adopt new technologies (eg, smart devices) [37], it is still necessary to collect more representative samples for future research. Considering the effects of social influences on technology adoption and use [36,38-40], it becomes more vital to collect samples with higher representativeness with regard to the socioeconomic status (SES). This is mainly because social influences are closely connected to educational levels and SES. That is, individuals with higher educational levels and SES are more likely to be affected by the subjective norms of their influential others who are more open to new technologies. Accordingly, it is recommended for future research to further consider the roles of educational levels and SES that are related to social influences by collecting more representative samples.

Finally, in terms of health information-seeking behaviors, we need to consider the following points. As Baumgartner and Hartmann [37] argued, searching for online health information is closely related to one's level of health anxiety. Moreover, research depending on information management theory [41-43] has stated that a person diagnosed with a chronic illness (eg, AIDS, cancer) will want to manage the amount of available information they are exposed to, rather than proactively seek information in order to reduce uncertainty. These findings commonly indicate that information-seeking behaviors through health apps are possibly moderated by people's actual health conditions. In other words, it is possible that people with chronic illnesses are less inclined to seek further information even though they have high levels of eHealth literacy as well as health-app use efficacy. Therefore, future research may consider further studying health-app use in relation to individuals' personal health conditions.

Conclusions

As a specific realm of mobile health, smartphone health apps are a significant form of technology that people have become increasingly interested in. However, we have had little understanding of the motivational factors that guide people to use health apps. Accordingly, this study aimed at exploring the effects of four cognitive factors—health consciousness, health information orientation, eHealth literacy, and health-app use efficacy—on the extent of health-app use. The results from a path analysis addressed the significant direct effect of health consciousness as well as strong mediating effects of health-app use efficacy. These findings contribute to broadening our comprehension of the new, digital dimensions of health management that revolve in particular around mobile technology.

Acknowledgments

This study was supported by the National Research Foundation of Korea Grant, funded by the Korean Government (NRF-2012S1A3A2033480) and Hallym University Research Fund (HRF-G-2012-2).

Conflicts of Interest

None declared.

References

1. Liu C, Zhu Q, Holroyd KA, Seng EK. Status and trends of mobile-health applications for iOS devices: A developer's perspective. Journal of Systems and Software 2011 Nov;84(11):2022-2033. [doi: 10.1016/j.jss.2011.06.049]
2. Fiordelli M, Diviani N, Schulz PJ. Mapping mHealth research: a decade of evolution. J Med Internet Res 2013 May;15(5):e95 [FREE Full text] [doi: 10.2196/jmir.2430] [Medline: 23697600]
3. Bender JL, Yue RY, To MJ, Deacken L, Jadad AR. A lot of action, but not in the right direction: systematic review and content analysis of smartphone applications for the prevention, detection, and management of cancer. J Med Internet Res 2013 Dec;15(12):e287 [FREE Full text] [doi: 10.2196/jmir.2661] [Medline: 24366061]
4. Rainie L. Pew Research Center. 2013. Cell phone ownership hits 91% of adults URL: http://www.pewresearch.org/fact-tank/2013/06/06/cell-phone-ownership-hits-91-of-adults/ [accessed 2014-05-07] [WebCite Cache ID 6POUgIeZx]
5. Fox S, Duggan M. Pew Research Center. 2012. Mobile health 2012 URL: http://www.pewinternet.org/files/old-media//Files/Reports/2012/PIP_MobileHealth2012_FINAL.pdf [accessed 2014-05-07] [WebCite Cache ID 6POV2piab]
6. Cummiskey M. There's an app for that. Smartphone use in health and physical education. J Phys Educ Recreat Danc 2011;82:24-30.
7. Kratzke C, Cox C. Smartphone technology and apps: Rapidly changing health promotion. Int Electron J Health Educ 2012;15:72-82.
8. Luxton DD, McCann RA, Bush NE, Mishkind MC, Reger GM. mHealth for mental health: Integrating smartphone technology in behavioral healthcare. Professional Psychology: Research and Practice 2011;42(6):505-512. [doi: 10.1037/a0024485]

9. Morris ME, Kathawala Q, Leen TK, Gorenstein EE, Guilak F, Labhard M, et al. Mobile therapy: case study evaluations of a cell phone application for emotional self-awareness. J Med Internet Res 2010 Apr;12(2):e10 [FREE Full text] [doi: 10.2196/jmir.1371] [Medline: 20439251]
10. Carter MC, Burley VJ, Nykjaer C, Cade JE. Adherence to a smartphone application for weight loss compared to website and paper diary: pilot randomized controlled trial. J Med Internet Res 2013 Apr;15(4):e32 [FREE Full text] [doi: 10.2196/jmir.2283] [Medline: 23587561]
11. Kirwan M, Vandelanotte C, Fenning A, Duncan MJ. Diabetes self-management smartphone application for adults with type 1 diabetes: randomized controlled trial. J Med Internet Res 2013 Nov;15(11):e235 [FREE Full text] [doi: 10.2196/jmir.2588] [Medline: 24225149]
12. 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 Aug;14(4):e111 [FREE Full text] [doi: 10.2196/jmir.2155] [Medline: 22868871]
13. Kamerow D. Regulating medical apps: which ones and how much? BMJ 2013 Oct 08;347(oct08 4):f6009-f6009. [doi: 10.1136/bmj.f6009]
14. Lim S, Xue L, Yen CC, Chang L, Chan HC, Tai BC, et al. A study on Singaporean women's acceptance of using mobile phones to seek health information. Int J Med Inform 2011 Dec;80(12):e189-e202. [doi: 10.1016/j.ijmedinf.2011.08.007] [Medline: 21956003]
15. Ministry of Science, ICT and Future Planning, Korea Internet and Security Agency. Seoul Korea. 2013. Korean Internet White Paper URL: http://isis.kisa.or.kr/eng/ebook/ebook.html [accessed 2014-05-07] [WebCite Cache ID 6POV5Rcof]
16. Kang I. IT Dong-A. 2013 Jul. Google "Please make mobile sites" URL: http://it.donga.com/15375/ [accessed 2014-05-07] [WebCite Cache ID 6POVMziAn]
17. Dutta-Bergman MJ. Primary sources of health information: comparisons in the domain of health attitudes, health cognitions, and health behaviors. Health Commun 2004;16(3):273-288. [doi: 10.1207/S15327027HC1603_1] [Medline: 15265751]
18. Norman CD, Skinner HA. eHEALS: The eHealth Literacy Scale. J Med Internet Res 2006;8(4):e27 [FREE Full text] [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
19. Basu A, Dutta MJ. The relationship between health information seeking and community participation: the roles of health information orientation and efficacy. Health Commun 2008 Jan;23(1):70-79. [doi: 10.1080/10410230701807121] [Medline: 18443994]
20. Michaelidou N, Hassan LM. The role of health consciousness, food safety concern and ethical identity on attitudes and intentions towards organic food. Int J Cons Stud 2008 Mar;32(2):163-170. [doi: 10.1111/j.1470-6431.2007.00619.x]
21. Dutta-Bergman MJ. The relation between health-orientation, provider-patient communication, and satisfaction: an individual-difference approach. Health Commun 2005;18(3):291-303. [doi: 10.1207/s15327027hc1803_6] [Medline: 16187933]
22. Moorman C, Matulich E. A model of consumers' preventive health behaviors, the role of health motivation and health ability. J Consum Res 1993;20:208-228.
23. Brodie M, Flournoy RE, Altman DE, Blendon RJ, Benson JM, Rosenbaum MD. Health information, the Internet, and the digital divide. Health Aff (Millwood) 2000;19(6):255-265 [FREE Full text] [Medline: 11192412]
24. Yun EK, Park HA. Consumers' disease information-seeking behaviour on the Internet in Korea. J Clin Nurs 2010 Oct;19(19-20):2860-2868. [doi: 10.1111/j.1365-2702.2009.03187.x] [Medline: 20598000]
25. Bandura A. Social functions of thought and action, a social cognitive theory. New Jersey: Prentice Hall, Inc; 1986.
26. Bodie GD, Dutta MJ. Understanding health literacy for strategic health marketing: eHealth literacy, health disparities, and the digital divide. Health Mark Q 2008 Jul;25(1-2):175-203. [doi: 10.1080/07359680802126301] [Medline: 18935884]
27. Fuchs C, Diamantopoulos A. Using single-item measures for construct measurement in management research. Conceptual issues and application guidelines. Die Betriebswirtschaft 2009;69:195-210.
28. Bergkvist L, Rossiter JR. The predictive validity of multiple-item versus single-item measures of the same constructs. Journal of Marketing Research 2007 May;44(2):175-184. [doi: 10.1509/jmkr.44.2.175]
29. Kline RB. Principles and Practice of Structural Equation Modeling. New York: Guilford Press; 2011.
30. Rattray J, Jones MC. Essential elements of questionnaire design and development. J Clin Nurs 2007 Feb;16(2):234-243. [doi: 10.1111/j.1365-2702.2006.01573.x] [Medline: 17239058]
31. Ferguson E, Cox T. Exploratory factor analysis: a user's guide. Int J Sel Assess 1993;1:84-94.
32. Lee HS, Lim JH. Structural equation modeling analysis and AMOS 6. Seoul, Korea: Bubmunsa Corporation; 2007.
33. Compeau DR, Higgins CA. Computer self-efficacy: Development of a measure and initial test. MIS Quarterly 1995 Jun;19(2):189-211. [doi: 10.2307/249688]
34. McCombs M. Mass media in the marketplace. J Monogr 1972;24:1-104.
35. Dutta-Bergman MJ. Complementarity in consumption of news types across traditional and new media. Journal of Broadcasting & Electronic Media 2004 Mar;48(1):41-60. [doi: 10.1207/s15506878jobem4801_3]
36. Katz E, Blumler JG, Gurevitch M. Uses and gratifications research. Public Opin Q 1973;37(4):509-523. [doi: 10.1086/268109]
37. Baumgartner SE, Hartmann T. The role of health anxiety in online health information search. Cyberpsychol Behav Soc Netw 2011 Oct;14(10):613-618. [doi: 10.1089/cyber.2010.0425] [Medline: 21548797]

38. Fulk J. Social construction of communication technology. Academy of Management Journal 1993 Oct 01;36(5):921-950. [doi: 10.2307/256641]
39. Fulk J, Schmitz J, Ryu D. Cognitive elements in the social construction of communication technology. Management Communication Quarterly 1995 Feb 01;8(3):259-288. [doi: 10.1177/0893318995008003001]
40. Zhou T. Understanding online community user participation: a social influence perspective. Internet Research 2011;21(1):67-81. [doi: 10.1108/10662241111104884]
41. Brashers DE. Communication and uncertainty management. J Commun 2001;51:477-497. [doi: 10.1093/joc/51.3.477]
42. Brashers DE. Information seeking and avoiding in health contexts. Human Communication Research 2002 Apr 01;28(2):258-271. [doi: 10.1093/hcr/28.2.258]
43. 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. Communication Monographs 2000 Mar;67(1):63-84. [doi: 10.1080/03637750009376495]

Abbreviations

CFI: comparative fit index
eHL: eHealth literacy
HAU: extent of health-app use
HAUE: health-app use efficacy
HC: health consciousness
HIO: health information orientation
IFI: incremental fit index
SEM: structural equation modeling
SES: socioeconomic status
SRMR: standardized root-mean squared residual

Edited by G Eysenbach; submitted 29.01.14; peer-reviewed by M Fiordelli, S Winter, C Rabin; comments to author 19.02.14; revised version received 26.03.14; accepted 28.04.14; published 09.05.14.

Please cite as:
Cho J, Park D, Lee HE
Cognitive Factors of Using Health Apps: Systematic Analysis of Relationships Among Health Consciousness, Health Information Orientation, eHealth Literacy, and Health App Use Efficacy
J Med Internet Res 2014;16(5):e125
URL: http://www.jmir.org/2014/5/e125/
doi:10.2196/jmir.3283
PMID:24824062

Letter to the Editor

Intervention Adherence is Related to Participant Retention: Implications for Research

John Alastair Cunningham[1,2], PhD

[1]Centre for Mental Health Research, The Australian National University, Canberra, Australia

[2]Centre for Addiction and Mental Health, Toronto, ON, Canada

Corresponding Author:
John Alastair Cunningham, PhD
Centre for Mental Health Research
The Australian National University
Building 63
Canberra, 0200
Australia
Phone: 61 02 6125 1859
Fax: 61 02 6125 0733
Email: john.cunningham@anu.edu.au

Related Article:

Comment on: http://www.jmir.org/2013/8/e162/

(J Med Internet Res 2014;16(5):e133) doi:10.2196/jmir.2948

KEYWORDS

research methods; randomized controlled trials; alcohol; Internet; brief intervention

Two challenging issues in Internet intervention research, as well as in other behavioral intervention trials, are ensuring that participants receive the intervention (adherence) and that their outcomes are captured at follow-up (retention) [1]. The interesting analysis presented by Murray et al [2] demonstrated that, at least in their study sample, the participant adherence and retention were positively related.

One issue to consider is whether this finding can be replicated in other study samples. It is possible that research involving, for example, different recruitment methods or with higher (or lower) retention rates, might not display this same positive relationship. To that purpose, results were examined from an Internet intervention trial that employed a proactive telephone recruitment method and obtained complete follow-up data for 86% of participants [3-5]. As with the Murray et al study [2], adherence (measured by the number of intervention participants logging onto a brief alcohol intervention, where N=92; 57 participants logged onto the intervention and 35 participants did not log on) and retention were strongly positively related (retention at 3-months: logged onto intervention=100%, did not log on=80%, $P<.001$; retention at 6-months: logged onto intervention=100%; did not log on=80%, $P<.001$; retention at 12-months: logged onto intervention=96%; did not log on=74.3%, $P=.002$; Fisher's Exact Tests).

Given that the positive relationship between adherence and retention can be replicated, what are the implications of this finding? From one perspective, the fact that these two key issues are related could underline the increased importance of obtaining a good retention rate. This is because the positive relationship of adherence to retention implies that a confound in the interpretation of the results is more likely as loss to follow-up (or reduced adherence to the intervention) increases. Alternatively, it could be argued that this positive relationship might reduce the importance of obtaining a good retention rate. This is because traditional intent-to-treat analysis assumes that participants who are lost to follow-up do not make any change in their behavior from baseline to follow-up (and are included as imputed values in the analysis based on this assumption). If it is then assumed that only those participants who accessed the intervention will actually make a change in their behavior, then the fact that participants who adhere to the intervention are more likely to follow-up can only increase the likelihood that participants who are lost to follow-up are less likely to have made a change in their behavior (thus validating the intent-to-treat analysis assumption). Determining which of these implications is correct is important, particularly in a field where low retention rates are an unfortunate reality in many research trials [1].

Conflicts of Interest

None declared.

References

1. 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]
2. Murray E, White IR, Varagunam M, Godfrey C, Khadjesari Z, McCambridge J. Attrition revisited: adherence and retention in a web-based alcohol trial. J Med Internet Res 2013;15(8):e162 [FREE Full text] [doi: 10.2196/jmir.2336] [Medline: 23996958]
3. Cunningham JA, Wild TC, Cordingley J, van Mierlo T, Humphreys K. A randomized controlled trial of an internet-based intervention for alcohol abusers. Addiction 2009 Dec;104(12):2023-2032 [FREE Full text] [doi: 10.1111/j.1360-0443.2009.02726.x] [Medline: 19922569]
4. Cunningham JA, Wild TC, Cordingley J, Van Mierlo T, Humphreys K. Twelve-month follow-up results from a randomized controlled trial of a brief personalized feedback intervention for problem drinkers. Alcohol 2010 Jun;45(3):258-262 [FREE Full text] [doi: 10.1093/alcalc/agq009] [Medline: 20150170]
5. Cunningham JA, Wild TC, Humphreys K. Who uses online interventions for problem drinkers? J Subst Abuse Treat 2011 Oct;41(3):261-264 [FREE Full text] [doi: 10.1016/j.jsat.2011.03.003] [Medline: 21632197]

Edited by G Eysenbach; submitted 09.09.13; peer-reviewed by E Murray, R Kok, S Langrial; accepted 15.05.14; published 21.05.14.

Please cite as:
Cunningham JA
Intervention Adherence is Related to Participant Retention: Implications for Research
J Med Internet Res 2014;16(5):e133
URL: http://www.jmir.org/2014/5/e133/
doi:10.2196/jmir.2948
PMID:24849941

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org